

**CARACTÉRISATION DES SYSTÈMES D'EXCLUSION DES PLASMIDES
CONJUGATIFS INCC**

par

Malika Humbert

Mémoire présenté au Département de biologie en vue
de l'obtention du grade de maître ès sciences (M.Sc.)

FACULTÉ DES SCIENCES
UNIVERSITÉ DE SHERBROOKE

Sherbrooke, Québec, Canada, septembre 2020

Le 14 septembre 2020

Le jury a accepté le mémoire de Malika Humbert dans sa version finale

Membres du jury

Professeur Vincent Burrus
Directeur de recherche
Département de biologie
Université de Sherbrooke

M. Benoît Leblanc
Évaluateur interne
Département de biologie
Université de Sherbrooke

Professeure Pascale Beauregard
Président-rapporteur
Département de biologie
Université de Sherbrooke

SOMMAIRE

La dissémination de résistance à de multiples antibiotiques est un enjeu majeur de santé publique. Ce phénomène est notamment dû à la conjugaison bactérienne, qui permet la transmission d'ADN entre deux bactéries présentes dans le même milieu, même entre espèces qui peuvent être phylogénétiquement très différentes. Les plasmides IncC sont des plasmides conjugatifs qui peuvent conférer une résistance à de multiples antibiotiques à leurs hôtes bactéries. Ils sont également capables de mobiliser des îlots génomiques lors de leur dissémination, tel que le *Salmonella* genomic island 1 (SGI1). Ces îlots participent eux aussi à la dissémination de gènes de résistance aux antibiotiques. L'objet de ce mémoire est de caractériser les systèmes d'exclusion des plasmides IncC, des mécanismes qui empêchent l'entrée d'un plasmide IncC dans une bactérie qui contient déjà un de ces plasmides. Pour cette caractérisation, une approche génétique a été privilégiée, consistant à observer l'effet de délétions de gènes sur l'efficacité de transferts de ces plasmides. Cette approche a été étayée par des analyses phylogénétiques, des prédictions de structures protéiques et une mutagénèse par insertion de transposon.

Les résultats décrits dans ce mémoire montrent l'existence de deux systèmes d'exclusion chez les plasmides IncC et les plasmides apparentés : un système d'exclusion d'entrée, impliquant EexC chez la cellule réceptrice et TraG chez la cellule donneuse ; et un système d'exclusion de surface, médié par SfxC chez la cellule réceptrice. Le système d'exclusion d'entrée peut être contourné par SGI1 lors de sa mobilisation, ce qui permet à cet îlot de se disséminer dans des populations bactériennes contenant un plasmide IncC. J'ai aussi montré que SGI1 exerce lui aussi une exclusion d'entrée : il s'agit du premier système d'exclusion d'entrée connu encodé par un îlot génomique mobilisable.

Mots-clés : conjugaison bactérienne, T4SS, plasmides IncC, SGI1, exclusion d'entrée, exclusion de surface.

REMERCIEMENTS

Tout d'abord, je tiens à remercier Vincent Burrus, mon directeur de recherche, pour m'avoir accueillie dans son laboratoire, et pour m'avoir encadrée et formée pendant plusieurs années. Je remercie également les membres actuels et passés du laboratoire, pour m'avoir fourni de nombreux conseils et tout particulièrement Kévin Huguet et Frédéric Coulombe, qui ont participé à l'obtention de résultats présentés ici. Je remercie également les membres de mon comité de conseillers, Benoît Leblanc et Pascale Beauregard, qui ont suivi ce projet depuis le début. Et enfin, je remercie mes amis et mes parents, qui m'ont soutenue moralement dans les moments plus difficiles.

TABLE DES MATIERES

SOMMAIRE.....	iv
REMERCIEMENTS.....	v
LISTE DES ABRÉVIATIONS	ix
LISTE DES TABLEAUX.....	x
LISTE DES FIGURES.....	xi
CHAPITRE 1 INTRODUCTION GÉNÉRALE	12
1.1 Mise en contexte.....	12
<i>1.1.1. Transfert conjugatif.....</i>	<i>12</i>
1.1.1.1 Conjugaison et plasticité du génome bactérien.....	12
1.1.1.2 Mécanisme général.....	14
1.1.1.3 Systèmes de sécrétion de type IV	17
<i>1.1.2. Exclusion.....</i>	<i>21</i>
1.1.2.1 Définition et importance.....	21
1.1.2.2 Exclusion d'entrée.....	23
1.1.2.3 Exclusion de surface.....	25
<i>1.1.3. Plasmides IncC.....</i>	<i>25</i>
1.1.3.1 Caractéristiques générales	25
1.1.3.2 Ilots génomiques mobilisés par les plasmides IncC.....	27
1.1.3.3 pVCR94 et SGI1 comme modèles d'études	28
1.1.3.4 Exclusion des plasmides IncC	31
1.2 Hypothèses	32
1.3 Objectifs	32
CHAPITRE 2 Exclusion d'entrée des plasmides conjugatifs des groupes d'incompatibilité IncA, IncC et des plasmides non typés proches	34
2.1 Introduction à l'article.....	34
<i>2.1.1. Originalité du travail et apport à la science.....</i>	<i>34</i>
<i>2.1.2. Apport véritable de chacun des auteurs.....</i>	<i>35</i>

2.1.3.	<i>Référence</i>	35
2.2	Article : Entry Exclusion of Conjugative Plasmids of the IncA, IncC, and Related Untyped Incompatibility Groups	36
2.2.1.	<i>Abstract</i>	36
2.2.2.	<i>Importance</i>	36
2.2.3.	<i>Keywords</i>	36
2.2.4.	<i>Introduction</i>	36
2.2.5.	<i>Results</i>	38
2.2.5.1	Identification of the gene mediating IncC plasmid entry exclusion in recipient cells	38
2.2.5.2	38
2.2.5.3	TraG is the entry exclusion factor in donor cells.	40
2.2.5.4	Clustering of EexC and TraG _C orthologues suggests multiple-entry exclusion groups	41
2.2.5.5	Confirmation of exclusion groups by mating assays.....	41
2.2.5.6	The variable C terminus of EexC defines the specificity of entry exclusion...43	
2.2.6.	<i>Discussion</i>	43
2.2.7.	<i>Materials and methods</i>	47
2.2.7.1	Bacterial strains and media.	47
2.2.7.2	Mating assays.....	47
2.2.7.3	Molecular biology methods.	48
2.2.7.4	Plasmid and strain construction.	48
2.2.7.5	Phylogenetic analyses.....	49
2.2.7.6	In silico typing of plasmids.....	49
2.2.7.7	Protein secondary structure prediction.	49
2.2.8.	<i>Supplemental material</i>	49
2.2.9.	<i>Acknowledgments</i>	49
2.2.10.	<i>References</i>	49
	CHAPITRE 3 Résultats complémentaires	52
3.1	Matériel et méthodes.....	52

3.1.1.	<i>Culture bactérienne</i>	52
3.1.2.	<i>Essais de conjugaison</i>	53
3.1.3.	<i>Méthodes de biologie moléculaire</i>	54
3.1.4.	<i>Construction des plasmides et des souches</i>	54
3.1.5.	<i>Mutagenèse par insertion de transposons</i>	56
3.1.6.	<i>Analyse phylogénétique</i>	56
3.1.7.	<i>Prédiction de structure secondaire de protéines</i>	57
3.2	Résultats	57
3.2.1.	<i>Exclusion d'entrée</i>	57
3.2.1.1	Chimères EexC-EexE	57
3.2.1.2	Exclusion d'entrée de SGI1	60
3.2.2.	<i>Exclusion de surface</i>	61
3.2.2.1	HDTM	61
3.2.2.2	Phylogénie	64
3.2.2.3	Prédiction de structure protéique	66
3.2.2.4	Essais d'exclusion avec Vcrx085.....	66
CHAPITRE 4 DISCUSSION GÉNÉRALE ET CONCLUSION		69
4.1	Exclusion d'entrée des plasmides IncC	69
4.2	Exclusion d'entrée de SGI1	71
4.3	Exclusion de surface de pVCR94	72
4.4	Conclusion et perspectives	74
ANNEXES		76
BIBLIOGRAPHIE		95

LISTE DES ABRÉVIATIONS

ADN	acide désoxyribonucléique
Ap	ampicilline
Cm	chloramphénicol
HDTM	mutagénèse par insertion de transposon
ICE	élément intégratif et conjugatif
IMC	complexe de la membrane interne
IncC	groupe d'incompatibilité C
kb	kilobase
Kn	kanamycine
MGI	îlot génomique mobilisable
MPS	stabilisation de la paire de conjugaison
Nx	acide nalidixique
OMC	complexe de la membrane externe
<i>oriT</i>	origine de transfert
Rf	rifampicine
SGI1	<i>Salmonella</i> genomic island I
Sm	streptomycine
Sp	spectinomycine
Su	sulfaméthoxazole
Tc	tétracycline
Tm	triméthoprim
ts	thermosensible
T4CP	protéine de couplage de type IV
T4SS	système de sécrétion de type IV

LISTE DES TABLEAUX

Tableau 1	Souches d' <i>E. coli</i> dérivées de K-12, plasmides et îlots génomiques utilisés	53
Tableau 2	Séquence nucléotidique des amorces utilisées	56
Tableau S1	Clusters de protéines Vcrx085 identiques	76

LISTE DES FIGURES

Figure 1	Mécanisme général de la conjugaison.....	15
Figure 2	Comparaison de la structure de T4SS de type P et de type F.....	18
Figure 3.	Comparaison des mécanismes d'incompatibilité et d'exclusion	22
Figure 4	Organisation génétique de pVCR94ΔX et SGI1ΔIn104.....	30
Figure 5	Effet des chimères EexC-EexE sur l'exclusion.....	59
Figure 6	Rôle de S009 dans l'exclusion de SGI1	61
Figure 7	Identification par HDTM d'un potentiel déterminant de l'exclusion	63
Figure 8	Analyse phylogénétique des ORFs homologues à <i>vcrx085</i>	65
Figure 9	Prédiction de la structure de Vcrx085	66
Figure 10	Rôle de <i>vcrx085</i> dans l'exclusion des plasmides IncC	68
Figure S1	Identification par HDTM d'un potentiel déterminant de l'exclusion, résultats complets.....	94

CHAPITRE 1

INTRODUCTION GÉNÉRALE

1.1 Mise en contexte

1.1.1. Transfert conjugatif

1.1.1.1 Conjugaison et plasticité du génome bactérien

Les bactéries sont des organismes présents presque partout sur notre planète, et capables de prospérer dans des écosystèmes extrêmement variés. Ce sont pourtant des êtres vivants qui peuvent sembler assez simples, en comparaison par exemple avec des eucaryotes multicellulaires. Cependant, cette simplicité s'accompagne d'une grande plasticité du génome, qui peut permettre aux populations bactériennes de s'adapter à de nouveaux milieux ou aux changements rapides de leur environnement. Cette évolution rapide du génome est due à des mutations ponctuelles, à des réarrangements chromosomiques par recombinaison ou au transfert horizontal, c'est-à-dire l'acquisition de nouvel ADN (acide désoxyribonucléique) en provenance du milieu extérieur ou d'autres bactéries (Patel, 2016).

Le transfert horizontal de gènes peut s'opérer par différents mécanismes. Les plus étudiés sont les suivants : la transformation, qui permet à une bactérie compétente de récupérer de l'ADN présent dans son milieu ; la transduction par des phages, pendant laquelle l'ADN transféré est encapsidé ; et la conjugaison, qui nécessite un contact direct entre la bactérie donneuse et la bactérie réceptrice (Soucy *et al.*, 2015).

La conjugaison bactérienne est médiée par des éléments conjugatifs, c'est-à-dire des molécules d'ADN qui encodent les protéines nécessaires à l'établissement d'un appareil conjugatif, au recrutement d'une cellule réceptrice et au transfert de leur propre ADN de la bactérie donneuse

à la réceptrice. Ces éléments génétiques mobiles peuvent être des molécules d'ADN séparées du génome, qui se répliquent et se maintiennent de manière autonome. On parle alors de plasmides conjugatifs. Les éléments intégratifs et conjugatifs (Integrative and Conjugative Elements, ou ICE), quant à eux, s'intègrent à un site spécifique du chromosome, sont répliqués et maintenus en même temps que celui-ci, et s'excisent avant de se transférer (Burrus, 2017).

Les éléments conjugatifs sont autotransmissibles car ils encodent la totalité des protéines nécessaires à leur transfert. Les éléments mobilisables, quant à eux, n'encodent pas toutes ces protéines, mais sont capables de se transférer en présence d'un élément conjugatif dans la même cellule. Ainsi, il existe des plasmides conjugatifs, des plasmides mobilisables, des éléments conjugatifs intégrés dans le chromosome (les ICEs) et des éléments mobilisables intégrés dans le chromosome (îlots génomiques mobilisables, ou MGI).

La conjugaison permet le transfert de molécules d'ADN qui peuvent dépasser 200 000 kilobases (kb) (Rozwandowicz *et al.*, 2018) entre deux bactéries qui peuvent être très éloignées phylogénétiquement. Certaines bactéries, telles qu'*Agrobacterium tumefaciens*, sont capables d'effectuer un transfert d'ADN vers une cellule eucaryote (Christie, 2004).

La capacité de la conjugaison à mobiliser de longues molécules d'ADN entre des bactéries potentiellement très différentes en fait un important facteur dans la dissémination de gènes de résistance aux antibiotiques. En effet, les éléments conjugatifs sont souvent porteurs d'intégrons ou de transposons, dans lesquels s'accumulent parfois de nombreux gènes d'antibiorésistance. Or la multirésistance aux antibiotiques est un phénomène qui a pris de plus en plus d'ampleur ces 30 dernières années et qui risque de s'aggraver. Un rapport sur la multirésistance aux antibiotiques, basé sur la situation en 2014, estime que cette année-là, environ 700 000 décès dans le monde étaient dus à des bactéries multirésistantes. Ce nombre pourrait s'élever à 10 millions de décès en 2050 (O'Neill, 2016). Ainsi, l'étude de la conjugaison bactérienne est essentielle pour mieux comprendre cette dissémination et, d'une manière plus globale, la plasticité du génome bactérien. Une compréhension plus précise de ce mécanisme pourrait

permettre de mieux prédire l'évolution des résistances aux antibiotiques, voire de produire des composés inhibiteurs de la conjugaison, pour en limiter les transferts.

1.1.1.2 Mécanisme général

Chez les bactéries à Gram négatif, le mécanisme de la conjugaison est remarquablement bien conservé (Guglielmini *et al.*, 2013). Il implique trois complexes protéiques : le relaxosome, impliqué dans la reconnaissance et la mobilisation de l'ADN à transférer ; le système de sécrétion de type IV (T4SS), aussi appelé pore de conjugaison ; et le pilus, qui permet de recruter la cellule réceptrice.

Les signaux à l'origine de l'initiation du transfert sont mal connus. Dans certaines conditions, l'activateur transcriptionnel permettant de réguler les gènes nécessaires à la conjugaison est exprimé plus fortement. Pour beaucoup d'éléments conjugatifs, cette régulation est mal comprise. Elle pourrait impliquer le quorum sensing, l'état physiologique de la cellule hôte (Grohmann *et al.*, 2018), ou encore la réponse SOS, comme c'est le cas chez les ICE SXT/R391 (Beaber *et al.*, 2004).

Lorsque le T4SS est formé, il permet la synthèse du pilus. Celui-ci est un polymère hélicoïdal composé majoritairement de sous-unités de piline et de quelques adhésines à son extrémité. Les monomères de piline sont ajoutés au niveau de la base du pilus, provoquant l'extension progressive de celui-ci.

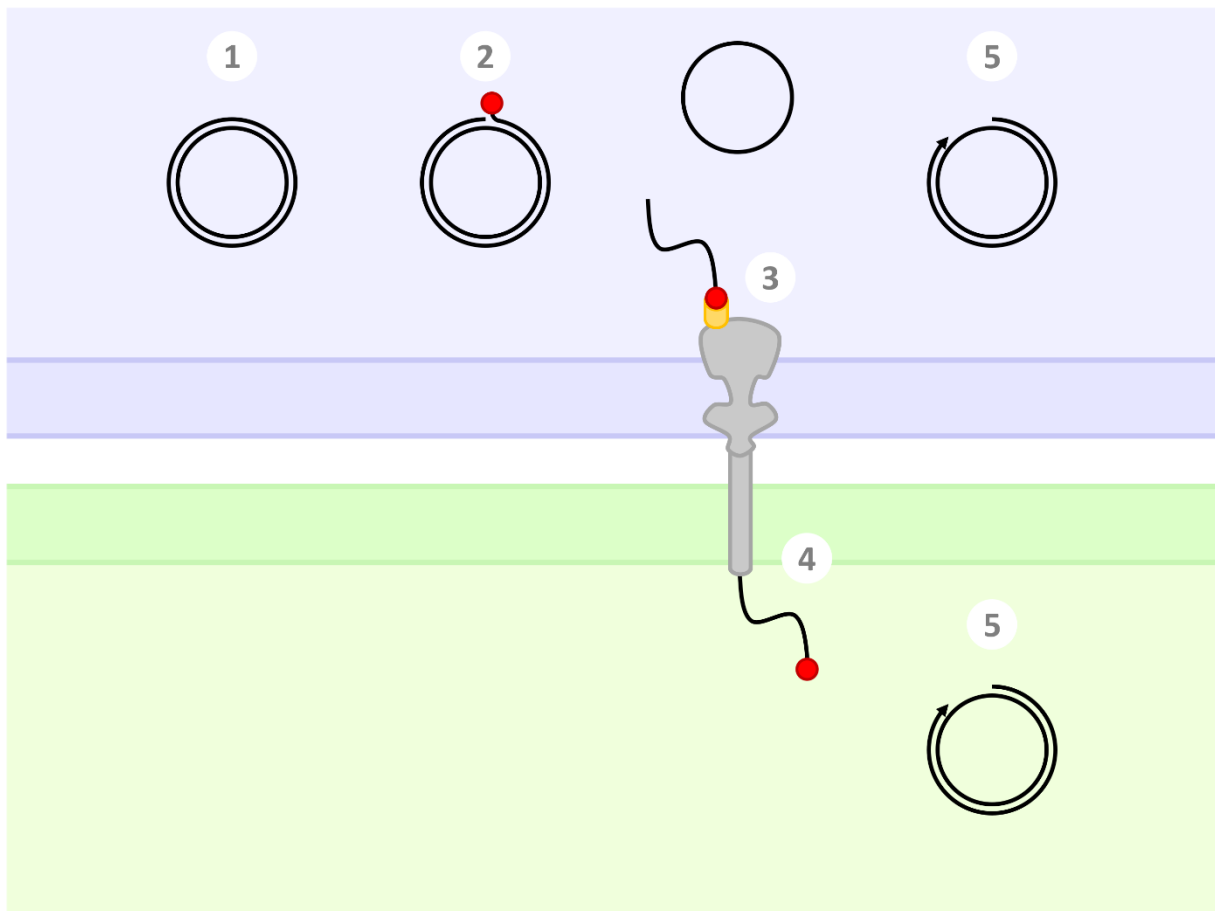


Figure 1. Mécanisme général de la conjugaison

(1) Plasmide présent dans la cellule donneuse, (2) clivage du site *nic* de l'*oriT* par la relaxase et formation du relaxosome (en rouge), (3) recrutement du relaxosome au niveau du T4SS par la T4CP (en jaune), (4) passage de l'ADN à travers le T4SS, (5) recircularisation et réplication.

L'initiation du transfert a lieu au niveau de l'origine de transfert (*oriT*), une séquence d'ADN contenant généralement plusieurs centaines de paires de bases (Cruz *et al.*, 2010). Celle-ci est reconnue par la relaxase, une endonucléase site-spécifique, qui permet le clivage simple brin de l'ADN au site *nic*. Le clivage forme une liaison covalente entre la relaxase et l'extrémité 5' du brin d'ADN clivé (Cabezón *et al.*, 2015; Grohmann *et al.*, 2018) (**figure 1** étape 2). Le complexe

formé par l'ADN, la relaxase et les protéines auxiliaires qui y sont liées est appelé relaxosome (Zechner *et al.*, 2012).

Le relaxosome est ensuite recruté au niveau du pore de conjugaison par la protéine de couplage de type IV (T4CP) (**figure 1** étape 3). Il s'agit d'une ATPase hexamérique possédant un canal central, qui interagit à la fois avec une protéine auxiliaire du relaxosome, et avec le pore de conjugaison. L'énergie produite par l'hydrolyse de l'ATP pourrait servir à pomper l'ADN à travers son canal central (Cabezón *et al.*, 2015).

Le T4SS permet le passage de la relaxase, sous forme dépliée, et de l'ADN simple brin à travers l'enveloppe de la cellule donneuse (**figure 1**, étape 4). Ce transfert nécessite de l'énergie, qui est fournie par des ATPases présentes du côté cytoplasmique de la membrane interne (Waksman, 2019). Il est difficile de savoir si l'ADN passe à travers le pilus ou non. Différentes observations directes de la conjugaison ont été effectuées. Dürrenberger *et al.* (1991) ainsi que Samuels *et al.* (2000) ont observé un contact direct entre la bactérie donneuse et la réceptrice, tandis que Babić *et al.* (2008) ont montré un transfert d'ADN par le pilus, entre deux cellules à distance l'une de l'autre. On ne sait pas de quelle manière l'ADN franchit l'enveloppe de la cellule réceptrice. À la fin du transfert, la relaxase permet la recircularisation du plasmide dans la cellule réceptrice, ainsi que l'initiation de sa réplication (**figure 1**, étape 5).

Le même mécanisme de transfert est retrouvé chez les ICE. Avant le transfert, l'ICE est excisé hors du chromosome. Un fois le transfert terminé, les deux copies de l'ICE sont intégrées dans le chromosome de la donneuse et de la réceptrice. Ces événements d'intégration et d'excision se font à un site spécifique du chromosome, appelé *attB*. L'intégrase permet la recombinaison site-spécifique du site *attP* de l'ICE dans le site *attB* du chromosome, ce qui mène à l'intégration de l'ICE, encadrés de deux sites nommés *attL* et *attR*. Avant un transfert conjugatif, l'excisionase est exprimée. Sa présence favorise la recombinaison entre *attL* et *attR*, elle aussi catalysée par l'intégrase, ce qui mène à l'excision de l'ICE et à la formation des sites *attP* et *attB* (Johnson et Grossman, 2015).

1.1.1.3 Systèmes de sécrétion de type IV

Les T4SS sont des complexes protéiques qui traversent la totalité de l'enveloppe bactérienne et sont retrouvés dans une grande variété de bactéries. Ce sont les seuls systèmes de sécrétions procaryotes capables de transporter aussi bien de l'ADN que des protéines (Waksman, 2019). Ils sont impliqués dans de nombreuses fonctions, telles que le transfert conjugatif, la sécrétion de molécules effectrices dans le cytoplasme de cellules eucaryotes, la contribution au développement d'un biofilm ou la sécrétion de toxines aux bactéries environnantes (Grohmann *et al.*, 2018).

Les T4SS possèdent donc une grande diversité fonctionnelle, qui s'accompagne également d'une grande diversité structurale. Chez les bactéries à Gram négatif, on en distingue trois types, initialement caractérisés par le pilus qui leur est associé : les T4SS de type P produisent des pili courts et rigides, ceux de type F des pili longs et flexibles, tandis que les T4SS de type I peuvent produire les deux sortes de pili (Christie, 2016). Pour chaque type de T4SS, un nombre restreint de systèmes a été caractérisé en détail. Ainsi, le système VirB/VirD4 d'*Agrobacterium tumefaciens* est le représentant le plus étudié du type P, le type F est principalement étudié chez le plasmide F de *Escherichia coli*, tandis que les principaux représentants du type I sont le plasmide R64 et le système Dot/Icm de *Legionella pneumophila* (Christie, 2016).

Les T4SS sont composés au minimum de douze protéines, et l'un des systèmes les plus simples est le système de type P VirB/VirD4 d'*A. tumefaciens*. VirB7, VirB9 et VirB10 forment le complexe ancré dans la membrane externe (OMC : outer membrane complex), tandis que le complexe de la membrane interne (IMC : inner membrane complex) est constitué de VirB3, VirB4, VirB6, VirB8 et d'une partie de VirB10. Le pilus est formé par VirB2 et VirB5, et les ATPases fournissant de l'énergie au système sont VirB4, VirB1 et VirD4, la protéine de couplage (**figure 2**).

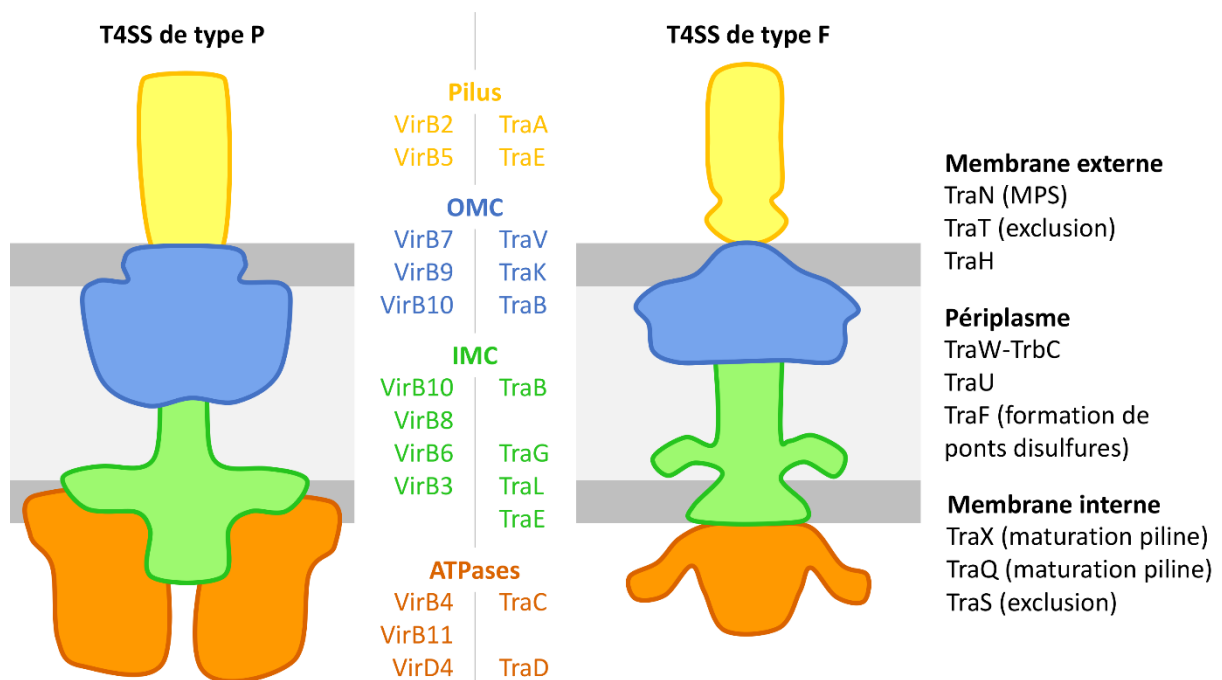


Figure 2. Comparaison de la structure de T4SS de type P et de type F

Protéines composant le pilus (en jaune), le complexe de la membrane externe (OMC, en bleu), le complexe de la membrane interne (IMC, en vert) et les ATPases (en orange) des T4SS de types P et F. Les noms des protéines homologues entre les deux T4SS sont écrits face à face. Les protéines du T4SS de type F ne possédant pas d'homologue dans le T4SS de type P sont listées à droite et leur fonction est indiquée entre parenthèses.

Les T4SS de type F contiennent bien plus de protéines que le système d'*A. tumefaciens*. Certaines d'entre elles sont homologues aux protéines VirB/VirD4 et leurs fonctions et leurs localisations ont souvent été déterminées d'après cette homologie (Lawley *et al.*, 2003). Une représentation simplifiée d'un T4SS de type F est présentée dans la **figure 2**.

Chez les T4SS de type F, seulement deux ATPases sont présentes : TraD, la protéine de couplage, et TraC. Cette dernière se trouve du côté cytoplasmique de la membrane interne, et

interagit avec le complexe de la membrane interne (Lawley *et al.*, 2003). Elle fournit l'énergie nécessaire à la formation du pilus ainsi qu'au transfert de l'ADN.

Le complexe de la membrane interne semble composé de TraE, TraL, TraG et d'une portion de TraB. Le complexe de la membrane externe, quant à lui, est composé de TraB, TraK et TraV. D'autres protéines, telles que TraN et TraH, sont situées dans la membrane externe, sans que l'on sache précisément si elles font partie du complexe de la membrane externe.

Les T4SS de type F contiennent un grand nombre de protéines qui ne présentent aucune homologie avec le système de type P VirB/VirD4, et dont la fonction est souvent moins bien caractérisée. Certaines d'entre elles, parfois appelées « auxiliaires », ne sont pas nécessaires pour la formation du T4SS, mais sont essentielles au transfert de l'ADN et à l'assemblage du pilus.

Le pilus de type F est formé par TraA. Initialement sous forme de propiline, celle-ci est intégrée dans la membrane interne, clivée, puis acétylée. Cette maturation de la propiline en piline nécessite TraQ et TraX, des protéines absentes du T4SS de type P. La piline est ensuite extraite de la membrane interne, un processus qui requiert l'hydrolyse d'ATP par TraC. L'extension du pilus se fait par ajout de monomères à la base de celui-ci, mais le détail de cette polymérisation est mal compris. Toutes les protéines nécessaires à la formation du T4SS, ainsi que TraF, TraH et TraW sont essentielles pour la formation du pilus (Lawley *et al.*, 2003).

Une fonction unique aux T4SS de type F est la stabilisation de la paire de conjugaison (mating pair stabilization, ou MPS), qui permet de renforcer la jonction entre les bactéries donneuse et réceptrice. Cette fonction est assurée par TraN, une adhésine ancrée dans la membrane externe, et par la partie C-terminale de TraG. TraN interagit avec TraV, présente dans le complexe de la membrane externe, et avec une protéine à la surface des bactéries réceptrices, qui pourrait être OmpA (Lawley *et al.*, 2003). Pour cette raison, TraN est suspectée de jouer un rôle dans la définition du spectre d'hôte du plasmide, car elle serait peut-être en mesure d'interagir de

manière spécifique avec les protéines de surface de certaines bactéries. TraG est une protéine aux multiples fonctions : la partie N-terminale, homologue à VirB6, fait partie du complexe de la membrane interne, et a un rôle structural dans le T4SS. La partie C-terminale n'a pas d'homologue dans le système VirB/VirD4, et est nécessaire à la stabilisation du contact entre donneuse et réceptrice (Lawley *et al.*, 2003). C'est aussi cette portion de la protéine qui est impliquée dans le mécanisme d'exclusion d'entrée (Audette *et al.*, 2007), qui sera discuté plus en détail dans la partie **1.1.2.2 Exclusion d'entrée (page 23)**.

Le T4SS de type F et le pilus qui lui est associé sont des structures dynamiques. En effet, en absence de réceptrice, le pilus s'allonge et se rétracte de manière cyclique, afin de scanner plus efficacement le milieu alentour. Une fois qu'il a établi le contact avec une cellule réceptrice, le pilus se rétracte, ce qui permet l'établissement d'un contact direct entre les deux cellules (Clarke *et al.*, 2008). Récemment, Hu *et al.* (2019) ont obtenu des images à haute résolution de T4SS de type F. Ils ont observé deux conformations majoritaires associées aux différentes fonctions du T4SS au cours du transfert. La première conformation correspond au T4SS lors de la phase d'extension et de rétraction du pilus. Cette conformation du T4SS n'est pas associée à la protéine de couplage et ne possède pas de canal central continu d'une membrane à l'autre. Lorsque la cellule réceptrice est contactée, le T4SS recrute TraD et le relaxosome, et subit d'importantes modifications de sa structure, principalement au niveau du complexe de la membrane externe et du cylindre central. Ces changements permettent la formation d'un canal qui s'étend de la membrane interne à la membrane externe. Dans la même étude, cette équipe a également observé des pili qui n'étaient attachés à aucun T4SS, et qui possédait parfois une plateforme basale, contenant bien moins de protéines que le T4SS en lui-même.

1.1.2. Exclusion

1.1.2.1 *Définition et importance*

Dès les premières études sur le plasmide F, il a été observé qu'une souche contenant ce plasmide pouvait difficilement en recevoir une deuxième copie (Lederberg *et al.*, 1952). Depuis, ce phénomène a été testé chez d'autres plasmides et il est très répandu : le transfert d'un plasmide vers une cellule contenant une copie de celui-ci est beaucoup moins efficace que vers une cellule vide (Garcillán-Barcia et de la Cruz, 2008).

Cette observation s'explique en fait par deux phénomènes distincts : l'incompatibilité et l'exclusion (**figure 3**). L'incompatibilité est une interférence entre les systèmes de maintenance des deux plasmides qui mène à la perte d'un des plasmides au fil des divisions cellulaires. Elle a lieu après le transfert conjugatif, typiquement lors de la réplication du génome et de la division de la bactérie hôte. L'exclusion, en revanche, agit lors du transfert et est médiée par des protéines n'ayant pas d'autre fonction que l'exclusion. En ce sens, on pourrait considérer l'incompatibilité comme un mécanisme « passif » et l'exclusion comme un mécanisme « actif ». Il existe deux mécanismes d'exclusion : on parle d'exclusion de surface lorsque le contact entre donneuse et réceptrice est inhibé, et d'exclusion d'entrée lorsque le transfert conjugatif est bloqué à une étape ultérieure. Certains éléments conjugatifs encode un seul de ces systèmes, d'autres possèdent les deux.

Garcillán-Barcia et de la Cruz (2008) ont consacré une revue de littérature à ce sujet. Ils ont trouvé des systèmes d'exclusion chez des ICE et des plasmides, autonomes et mobilisables, aussi bien retrouvés dans des bactéries à Gram négatif, à Gram positif, ou dans des archées. La présence d'un même mécanisme dans des éléments aussi variés suggère qu'il s'agit d'une fonction importante, voire essentielle. D'après ces auteurs, l'exclusion pourrait être présente chez tous les plasmides conjugatifs, et permettrait au plasmide présent dans la réceptrice d'éviter d'être perdu suite à l'entrée d'un plasmide incompatible dans la même cellule. En évitant les

transferts entre deux bactéries possédant le même plasmide, elle pourrait aussi réduire la consommation d'énergie des bactéries hôtes, et augmenter la disponibilité des donneuses pour leur permettre de transmettre le plasmide à des réceptrices vides. D'après Gago-Córdoba *et al.* (2019) l'exclusion permettrait aussi d'éviter une situation de transfert bidirectionnel entre deux donneuses, qui pourrait mener à une interruption du transfert, voire à la lyse des cellules.

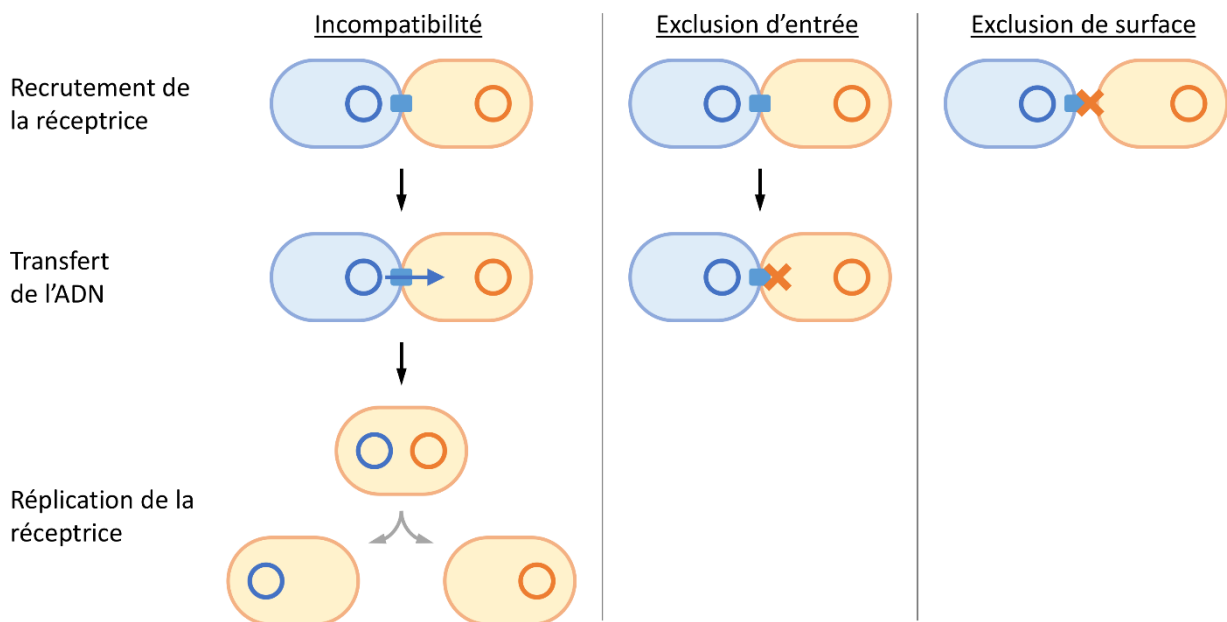


Figure 3. Comparaison des mécanismes d'incompatibilité et d'exclusion

Représentation de différentes étapes d'un transfert conjugatif en présence d'un mécanisme d'incompatibilité, d'exclusion d'entrée ou d'exclusion de surface.

L'exclusion a été étudiée chez les plasmides IncF (Audette *et al.*, 2007), IncP (Haase *et al.*, 1996), IncI (Sakuma *et al.*, 2013), IncH (Gunton *et al.*, 2008), IncN/IncW (Pohlman *et al.*, 1994) et IncC (Carraro *et al.*, 2017a) ainsi que chez les ICE SXT/R391 (Marrero et Waldor, 2005). Elle a aussi été étudiée dans des éléments de bactéries à Gram positif, tels que ICEBs1 (Dunny et Hirt, 2019) ou pLS20 (Gago-Córdoba *et al.*, 2019) de *Bacillus subtilis*. D'une manière générale, les facteurs d'exclusion sont assez différents d'un système à l'autre, et ne partagent pas d'homologie entre eux.

1.1.2.2 Exclusion d'entrée

Chez le plasmide F, l'exclusion d'entrée est médiée par TraS, une protéine de la membrane interne, encodée par *traS*, un gène situé directement en aval de *traG* (Jalajakumari *et al.*, 1987). Lorsque TraS est surexprimée dans la réceptrice, le transfert du plasmide F est réduit d'un facteur 1000 environ (Audette *et al.*, 2007). Cette exclusion nécessite la reconnaissance de TraG, présent dans le T4SS de la donneuse, par TraS, présent dans la membrane interne de la réceptrice (Anthony *et al.*, 1999). Le plasmide R100, proche du plasmide F et appartenant à la famille IncFII, encode une protéine TraS fonctionnelle, homologue à celle encodée par le plasmide F, et induisant une exclusion d'entrée. Cependant, les plasmides F et R100 ne s'excluent pas mutuellement (Anthony *et al.*, 1999). Ces deux plasmides ont ainsi été classés dans deux groupes d'exclusion distincts. Audette *et al.* (2007) ont étudié ce qui permettait la spécificité de cette exclusion : il s'agit de la reconnaissance par TraS de la portion centrale de TraG. En effet, lorsque l'on échange les portions centrales de TraG_F et TraG_{R100}, on change leur groupe d'exclusion.

Les ICE de la famille SXT/R391 possèdent aussi un système d'exclusion d'entrée. Le facteur d'exclusion est appelé EexS chez SXT et EexR chez R391 (Marrero et Waldor, 2005). Comme TraS_F, ces protéines sont ancrées dans la membrane interne, et leur présence dans la cellule réceptrice suffit à empêcher le transfert de l'ICE. SXT et R391 appartiennent à deux groupes d'exclusion différents : EexS exclue SXT mais pas R391, et EexR exclue R391 mais pas SXT. Tous les ICE de cette famille font partie d'un de ces deux groupes d'exclusion (Marrero et Waldor, 2007a). Comme pour le plasmide F, c'est TraG qui est reconnu par Eex et cette reconnaissance permet la spécificité de l'exclusion. Plus précisément, l'échange de trois acides aminés (606 à 608) de TraG_S et TraG_R suffit à intervertir leurs groupes d'exclusion (Marrero et Waldor, 2005). De même, les parties C-terminales de EexS et EexR sont responsables de la reconnaissance de TraG_S ou TraG_R (Marrero et Waldor, 2007b).

Le système d'exclusion des plasmides IncI a été étudié chez les plasmides R144, R64 et R621a. L'exclusion d'entrée est médiée par ExcA, une protéine de la membrane interne (Hartskeerl *et al.*, 1986). Sa présence dans la réceptrice, mais pas dans la donneuse, est nécessaire pour l'exclusion d'entrée (Sakuma *et al.*, 2013). Il existe aussi différents groupes d'exclusion chez les plasmides IncI, ce qui est dû à la reconnaissance par ExcA de TraY, un homologue fonctionnel de TraG_F. ExcA ne reconnaît pas la même portion de TraY chez tous les plasmides : chez R64, il s'agit de la partie centrale, tandis que c'est la partie C-terminale qui est reconnue chez R621a (Sakuma *et al.*, 2013).

Il existe plusieurs similitudes entre ces différents systèmes d'exclusion. Premièrement, tous ces systèmes impliquent un homologue de TraG_F dans la cellule donneuse et une protéine d'exclusion dans la réceptrice. Ces protéines d'exclusion ne sont pas homologues d'une famille de plasmides à une autre, mais ce sont généralement des protéines d'une centaine d'acides aminés, ancrées dans la membrane interne et encodées par un gène situé à proximité de l'homologue de *traG*. La présence de ce facteur d'exclusion dans la réceptrice est suffisante pour induire une exclusion d'entrée, même si le reste du plasmide n'est pas présent.

Ces différents systèmes d'exclusion comportent plusieurs groupes d'exclusion, dus à la spécificité entre le facteur d'exclusion et l'homologue de TraG_F. Dans plusieurs systèmes, les portions de ces deux protéines qui induisent cette spécificité sont cytoplasmiques (Audette *et al.*, 2007; Marrero et Waldor, 2007b). Il est difficile d'expliquer comment deux protéines dans le cytoplasme de deux bactéries différentes peuvent interagir. Pour le plasmide F, les essais de double hybride, de cross-link et d'immunoprécipitation pour détecter l'interaction entre TraG et TraS ont été infructueux (Audette *et al.*, 2007). De même, le mécanisme empêchant le transfert de l'ADN n'a pas été élucidé. Cependant, une quantification de la dégradation de l'ADN dans des mini-cellules a montré que l'exclusion d'entrée n'augmentait pas la dégradation de l'ADN transféré dans la réceptrice (Sheehy *et al.*, 1972). Il est donc probable que l'exclusion d'entrée joue un rôle lors du passage de l'ADN dans le T4SS.

Chez les bactéries à Gram négatif, le seul système d'exclusion qui semble fonctionner d'une manière totalement différente est celui des plasmides IncH. En effet, le plasmide R27 encode deux protéines d'exclusion : EexA, responsable de l'exclusion d'entrée, et EexB, responsable de l'exclusion de surface. Pour que l'exclusion ait lieu, il faut que ces protéines soient présentes à la fois dans la donneuse et dans la réceptrice (Gunton *et al.*, 2008). Ce mécanisme ne semble pas impliquer d'homologue de TraG_F.

1.1.2.3 *Exclusion de surface*

Chez les bactéries à Gram négatif, le seul système d'exclusion de surface qui a été bien caractérisé est celui du plasmide F. Il implique TraT, une protéine de la membrane externe, dont la présence réduit fortement la formation de paires de conjugaison stables (Achtman *et al.*, 1977). Sa surexpression dans une cellule réceptrice dénuée du plasmide F est suffisante pour empêcher le transfert de celui-ci. TraT semble aussi jouer un rôle dans la virulence de la bactérie, plus précisément en augmentant sa résistance au système immunitaire de son hôte. Moll *et al.* (1980) suggèrent que ceci pourrait être dû à la présence en grande quantité de TraT à la surface de la cellule, qui pourrait limiter la reconnaissance de certains antigènes par le système immunitaire. Le mécanisme permettant l'exclusion de surface par TraT est lui aussi mal compris, mais il est possible que la présence de TraT en grande quantité empêche la reconnaissance de la bactérie réceptrice par la bactérie donneuse, et par conséquent l'établissement d'une paire de conjugaison stable entre ces deux bactéries.

1.1.3. Plasmides IncC

1.1.3.1 *Caractéristiques générales*

Les plasmides du groupe d'incompatibilité C (IncC) sont des plasmides conjugatifs retrouvés chez de nombreuses Entérobactéries (notamment dans les genres *Escherichia*, *Klebsiella*, *Salmonella*, *Enterobacter*, *Aeromonas*) et Vibrionacées, tel que *Vibrio cholerae* (Harmer et

Hall, 2015). Ils peuvent conférer un grand nombre de résistances aux antibiotiques d'intérêt clinique tels que les carbapénèmes ou les céphalosporines de troisième génération (Harmer et Hall, 2015). Ils ont été isolés chez l'humain, dans des milieux hospitaliers, chez des animaux d'élevage et dans l'environnement, et ce sur tous les continents. Puisqu'ils ont un spectre d'hôte assez large, sont présents dans des environnements variés et peuvent conférer des multirésistances aux antibiotiques, ils sont des acteurs majeurs de la dissémination d'antibiorésistance.

Les plasmides IncC ont longtemps été regroupés avec les plasmides IncA sous le terme de groupe d'incompatibilité IncA/C. Le premier représentant des IncA est pRA1, et a été isolé chez *Aeromonas hydrophila* en 1971, et depuis 2016, d'autres plasmides IncA ont été identifiés. Cependant, Ambrose *et al.* (2018) ont montré que le regroupement des plasmides IncA et IncC était erroné. Malgré leur appartenance à des groupes d'incompatibilité différents, les plasmides IncC et IncA présentent de fortes ressemblances, notamment au niveau de l'organisation génétique et des gènes essentiels à la réplication, à la partition, et au transfert.

Les plasmides IncC sont des plasmides autonomes, autant pour leur réplication que pour leur transfert. Ils possèdent donc des gènes impliqués dans le transfert conjugatif, la réplication, la partition, le métabolisme de l'ADN et la régulation, ainsi qu'un système toxine/antitoxine et de nombreux gènes de fonction inconnue (Carraro *et al.*, 2014a; Hancock *et al.*, 2017; Harmer et Hall, 2015). Ils encodent un T4SS de type F et la fonction de la plupart des gènes impliqués dans le transfert a été déduite par homologie avec ceux du plasmide F et d'autres plasmides semblables.

Les gènes impliqués dans le transfert conjugatif sont pour la plupart sous le contrôle d'un régulateur transcriptionnel appelé AcaCD (Carraro *et al.*, 2014a). AcaCD est un hétéromère formé des protéines AcaC et AcaD, homologues à FlhC et FlhD, impliqués dans la régulation de la synthèse du flagelle. Carraro *et al.* (2014a) ont identifié les sites de liaison d'AcaCD aux plasmides IncC : ce régulateur se fixe à dix-sept promoteurs, notamment en amont de gènes

nécessaires à la mobilisation de l'ADN et à la formation du pore de conjugaison. Récemment, Hancock *et al.* (2020) ont identifié un second activateur transcriptionnel, nommé AcaB, lui aussi activé par AcaCD. AcaB se fixe au promoteur en amont d'*acaCD* et en active la transcription, ce qui mène à une boucle d'activation d'AcaB et AcaCD. Deux répresseurs transcriptionnels, Acr1 et Acr2, permettent la répression du promoteur en amont d'*acaCD* (Carraro *et al.*, 2014a). La manière dont Acr1 et Acr2 entrent en compétition avec AcaB au niveau de ce promoteur n'a pas encore été élucidée, de même que le signal initial permettant le déclenchement de la boucle d'activation AcaB/AcaCD.

1.1.3.2 Ilots génomiques mobilisés par les plasmides IncC

Le motif reconnu par AcaCD est aussi retrouvé dans différentes familles d'îlots génomiques, intégrés dans le chromosome de bactéries telles que *Salmonella enterica*, *Shewanella putrefaciens* et plusieurs espèces de *Vibrio* (Carraro *et al.*, 2014a, 2015, 2016). Il a été vérifié expérimentalement que les plasmides IncC pouvaient mobiliser certains de ces îlots, qui portent parfois des gènes avantageux pour leur hôte, tels que des gènes de virulence ou de résistance aux antibiotiques ou des systèmes de restriction/modification pouvant protéger leur hôte d'infections virales (Carraro *et al.*, 2015).

SGI1 (*Salmonella* genomic island 1) et les îlots qui lui sont apparentés sont retrouvés principalement chez *S. enterica* et *Proteus mirabilis*. SGI1 contient l'intégron In104, porteur de nombreuses résistances aux antibiotiques. En présence d'un plasmide IncC ou IncA, SGI1 peut se répliquer (Huguet *et al.*, 2020) et transférer à haute fréquence : son transfert est plus efficace que celui du plasmide mobilisateur. De plus, le transfert du plasmide est beaucoup moins efficace en présence de SGI1 que lorsqu'il est seul, et le taux de co-transfert est plus bas qu'attendu (Carraro *et al.*, 2014a; Harmer *et al.*, 2016). Cela est expliqué en partie par le fait qu'en présence d'AcaCD, SGI1 produit trois protéines, TraGs, TraHs et TraNs, qui s'incorpore dans le T4SS produit par le plasmide IncC et y remplace TraGc, TraHc et TraNc (Carraro *et al.*, 2017a). De plus, lorsqu'une souche contenant un plasmide IncC et SGI1 est cultivée sans

sélection par des antibiotiques, le plasmide est progressivement perdu (Harmer *et al.*, 2016). Cette déstabilisation des plasmides IncC pourrait permettre à SGI1 d'augmenter sa propre stabilité, car la présence d'AcaCD dans la cellule entraîne l'activation du gène *xis*, qui provoque l'excision de l'îlot, une forme moins stable que l'îlot intégré dans le chromosome. SGI1 ne se contente donc pas d'emprunter le système de conjugaison des plasmides IncC, mais est plutôt un parasite qui modifie ce système à son avantage, tout en déstabilisant ces plasmides.

Les plasmides IncC mobilisent d'autres îlots génomiques, tels que MGIV*mi*1, qui a été isolé chez *Vibrio mimicus* (Carraro *et al.*, 2014a), et MGIV*ch*Hai6, qui provient d'une souche clinique de *V. cholerae* (Carraro *et al.*, 2016). Comme c'est le cas pour SGI1, l'excision de MGIV*mi*1 et de MGIV*ch*Hai6 est induite lors de l'expression d'AcaCD (Carraro *et al.*, 2014a, 2016). Ces îlots n'encodent pas de protéines impliquées dans le T4SS. Un gène homologue à *mobI*, qui fait partie du relaxosome des plasmides IncC, et permet la reconnaissance de leur *oriT*, est retrouvé en aval d'un site putatif de liaison d'AcaCD (Rivard *et al.*, 2020). Ces îlots modifient donc le relaxosome des plasmides IncC afin qu'il puisse reconnaître leur *oriT*.

1.1.3.3 *pVCR94 et SGI1 comme modèles d'études*

Le plasmide IncC servant de modèle d'étude principal pour ce projet est pVCR94. Celui-ci était présent dans une souche pandémique de *V. cholerae* O1 El Tor isolée lors d'une épidémie de choléra dans un camp de réfugiés rwandais en République démocratique du Congo en 1994. Ce plasmide a déjà été bien caractérisé (Carraro *et al.*, 2014b, 2014a). Il peut transférer à assez haute fréquence ($>10^{-2}$ transconjugants/donneuses) chez *V. cholerae* et *E. coli*. Son *oriT* est située entre les gènes *vcrl52* et *mobI* (Carraro *et al.*, 2014b; Hegyi *et al.*, 2017). Il est porteur de gènes de résistance au co-trimoxazole, à l'ampicilline, à la streptomycine, au chloramphénicol et à la tétracycline. Ainsi, pour en simplifier l'utilisation en laboratoire, une grande région de ce plasmide, contenant principalement des transposons et des intégrons, a été enlevée. Le plasmide ainsi réduit, pVCR94ΔX (Genbank KF551948, **figure 4**), transfère à la même fréquence que pVCR94 (Carraro *et al.*, 2014b). pVCR94ΔX porte un gène de résistance

au sulfaméthoxazole, et trois autres versions de ce plasmide ont été construites en insérant de nouveaux gènes de résistance : pVCR94ΔX2, pVCR94ΔX3 et pVCR94ΔX4, qui sont résistants à la spectinomycine, à la kanamycine et au chloramphénicol, respectivement. Ces constructions ont grandement facilité l'étude de l'interaction de pVCR94 avec d'autres éléments génétiques mobiles (Carraro *et al.*, 2016, 2017b, 2017a) ainsi que la construction de nombreux mutants de délétions (Carraro *et al.*, 2014b, 2014a, 2017a). Pour faciliter la compréhension et simplifier l'écriture, les dérivés de pVCR94 seront appelés pVCR94^{Sp}, pVCR94^{Kn} et pVCR94^{Cm} dans la suite de ce mémoire.

Dans ce projet, SGI1 sera aussi étudié, car sa capacité à remplacer certaines protéines dans le T4SS des plasmides IncC lui permet de contourner leur exclusion d'entrée (Carraro *et al.*, 2017a). Il contient l'intégron In104, qui porte de nombreux gènes de résistance à des antibiotiques. Pour en simplifier l'utilisation en laboratoire, cet intégron a été enlevé et remplacé par une cassette de résistance à la kanamycine (Carraro *et al.*, 2017a). Cette construction est appelée SGI1ΔIn104 (**figure 4**). Pour ce mémoire, j'ai également construit une version de SGI1 dans laquelle l'intégron In104 est remplacé par un gène de résistance au chloramphénicol. Pour distinguer facilement ces deux versions, elles seront par la suite nommées SGI1^{Kn} et SGI1^{Cm}.

1.1.3.4 Exclusion des plasmides IncC

C'est en étudiant l'interaction entre SGI1 et pVCR94 que notre laboratoire a observé l'existence d'un système d'exclusion chez les plasmides IncC (Carraro *et al.*, 2017a). En effet, afin de favoriser son propre transfert, SGI1 remplace trois protéines du T4SS de pVCR94, TraG_C, TraH_C et TraN_C, par les protéines homologues TraG_S, TraH_S et TraN_S. Ce remplacement a pour avantage d'augmenter la fréquence de transfert de SGI1 au détriment de celle de pVCR94. Par ailleurs, le remplacement de TraG_C par TraG_S permet à SGI1 de contourner l'exclusion exercée par pVCR94.

Un autre groupe de recherche a observé ce phénomène d'exclusion chez les plasmides IncC. Ambrose *et al.* (2018) ont cherché à démêler les relations entre les plasmides IncC et IncA, qui ont longtemps été classés ensemble dans le groupe d'incompatibilité IncA/C. Cependant, ce regroupement n'était pas basé sur une véritable étude de l'incompatibilité entre ces plasmides, mais sur l'observation que ces plasmides ne sont jamais retrouvés dans la même souche, et qu'ils possèdent des réplicons similaires. Ambrose *et al.* ont montré que ces plasmides peuvent en réalité se maintenir dans une même souche, et que l'impossibilité d'introduire ces plasmides dans une même bactérie est due à un phénomène d'exclusion et non d'incompatibilité.

1.2 Hypothèses

Ainsi, mon projet de maîtrise s'est construit autour des hypothèses suivantes :

- les plasmides IncC encodent au moins un facteur d'exclusion, responsable d'une exclusion d'entrée ou de surface.
- SGI1 peut contourner ce système d'exclusion en modifiant le T4SS du plasmide qui le mobilise.
- TraG est impliqué dans ce système d'exclusion.

1.3 Objectifs

Bien que l'existence d'un système d'exclusion chez les plasmides IncC ait été confirmée, aucune étude préalable à mon projet de maîtrise n'avait identifié le ou les facteurs responsables de cette exclusion. L'objectif principal de ma maîtrise a donc été de caractériser l'exclusion médiée par les plasmides conjugatifs IncA et IncC.

Plus spécifiquement, mes travaux ont consisté à :

- identifier le ou les facteurs d'exclusion des plasmides IncA et IncC
- tenter de définir le rôle de TraG dans l'exclusion
- déterminer s'il existe un ou plusieurs groupes d'exclusion chez les plasmides IncA et IncC et le cas échéant, comprendre ce qui définit la spécificité de ces groupes d'exclusion.

- comprendre comment SGI1 interagit avec le mécanisme d'exclusion.
- puisque SGI1 encode une protéine TraG, et que celles-ci sont souvent impliquées dans l'exclusion, déterminer si SGI1 possède lui aussi un système d'exclusion.

CHAPITRE 2 EXCLUSION D'ENTRÉE DES PLASMIDES CONJUGATIFS DES GROUPE D'INCOMPATIBILITÉ INCA, INCC ET DES PLASMIDES NON TYPÉS PROCHES

2.1 Introduction à l'article

2.1.1. Originalité du travail et apport à la science

Dans l'article qui suit, nous avons étudié le système d'exclusion d'entrée des plasmides IncC, d'importants vecteurs de dissémination de multirésistance aux antibiotiques. Bien que l'existence d'un système d'exclusion ait déjà été signalée par deux articles précédents (Ambrose *et al.*, 2018; Carraro *et al.*, 2017a), ce système n'avait pas encore été caractérisé. Les principales conclusions de cette étude sont résumées ci-dessous :

- Nous avons pu identifier le gène responsable de l'exclusion d'entrée chez les plasmides IncC. Situé en aval du gène TraG, *vcx145* encode une petite protéine de 120 acides aminés, probablement ancrée dans la membrane interne. Ce gène a été renommé *eexC* (entry exclusion of IncC plasmids). Il semble que EexC ne soit pas le seul facteur d'exclusion chez les plasmides IncC, mais aucun autre facteur d'exclusion n'a été identifié jusqu'à présent.
- EexC reconnaît TraG dans le pore de conjugaison de la donneuse, et c'est cette exclusion médiée par EexC que SGI1 est capable de contourner en remplaçant TraG dans le T4SS.
- Au sein des plasmides qui encodent un homologue de EexC, trois groupes d'exclusion distincts ont pu être identifiés par phylogénie et confirmés expérimentalement. Ainsi, EexC est retrouvé chez les plasmides IncC et IncA ; EexA est encodé par des plasmides

d'*Aeromonas salmonicida*, tels que pAsa4c ; tandis que EexB provient de pAhD4-1, un plasmide isolé chez *A. hydrophila*.

- La spécificité de Eex envers TraG, qui permet la reconnaissance de ces groupes d'exclusion, est conférée par la partie C-terminale de Eex.

2.1.2. Apport véritable de chacun des auteurs

J'ai développé l'approche expérimentale avec Vincent Burrus. J'ai effectué les manipulations nécessaires pour les figures 1, 2, 3A, 3B et 4, avec l'aide de Frédéric Coulombe pour les figures 2 et 4C. Kévin T. Huguet a obtenu les résultats de la figure 3C. Les constructions génétiques ont été fournies par Kévin T. Huguet et moi, avec le support de Frédéric Coulombe. J'ai préparé les figures sous la supervision de Vincent Burrus. Le manuscrit a été rédigé par Vincent Burrus et moi. Tous les auteurs ont contribué à la relecture et à la correction du manuscrit.

2.1.3. Référence

Humbert M, Huguet KT, Coulombe F, Burrus V. 2019. Entry exclusion of conjugative plasmids of the IncA, IncC, and related untyped incompatibility groups. J Bacteriol 201:e00731-18. <https://doi.org/10.1128/JB.00731-18>



Entry Exclusion of Conjugative Plasmids of the IncA, IncC, and Related Untyped Incompatibility Groups

Malika Humbert,^a Kévin T. Huguet,^a Frédéric Coulombe,^a Vincent Burrus^a

^aDépartement de biologie, Université de Sherbrooke, Sherbrooke, Quebec, Canada

ABSTRACT Conjugative plasmids of incompatibility group C (IncC), formerly known as A/C₂, disseminate antibiotic resistance genes globally in diverse pathogenic species of *Gammaproteobacteria*. *Salmonella* genomic island 1 (SGI1) can be mobilized by IncC plasmids and was recently shown to reshape the conjugative type IV secretion system (T4SS) encoded by these plasmids to evade entry exclusion. Entry exclusion blocks DNA translocation between cells containing identical or highly similar plasmids. Here, we report that the protein encoded by the entry exclusion gene of IncC plasmids (*eexC*) mediates entry exclusion in recipient cells through recognition of the IncC-encoded TraG_C protein in donor cells. Phylogenetic analyses based on EexC and TraG_C homologs predicted the existence of at least three different exclusion groups among IncC-related conjugative plasmids. Mating assays using Eex proteins encoded by representative IncC and IncA (former A/C₁) and related untyped plasmids confirmed these predictions and showed that the IncC and IncA plasmids belong to the C exclusion group, thereby explaining their apparent incompatibility despite their compatible replicons. Representatives of the two other exclusion groups (D and E) are untyped conjugative plasmids found in *Aeromonas* sp. Finally, we determined through domain swapping that the carboxyl terminus of the EexC and EexE proteins controls the specificity of these exclusion groups. Together, these results unravel the role of entry exclusion in the apparent incompatibility between IncA and IncC plasmids while shedding light on the importance of the TraG subunit substitution used by SGI1 to evade entry exclusion.

IMPORTANCE IncA and IncC conjugative plasmids drive antibiotic resistance dissemination among several pathogenic species of *Gammaproteobacteria* due to the diversity of drug resistance genes that they carry and their ability to mobilize antibiotic resistance-conferring genomic islands such as SGI1 of *Salmonella enterica*. While historically grouped as “IncA/C,” IncA and IncC replicons were recently confirmed to be compatible and to abolish each other’s entry into the cell in which they reside during conjugative transfer. The significance of our study is in identifying an entry exclusion system that is shared by IncA and IncC plasmids. It impedes DNA transfer to recipient cells bearing a plasmid of either incompatibility group. The entry exclusion protein of this system is unrelated to any other known entry exclusion proteins.

KEYWORDS antibiotics, conjugation, conjugative plasmid, entry exclusion, IncA, IncA/C, IncC, resistance, T4SS, VirB6

Conjugative plasmids of incompatibility group C (IncC) are large, broad-host-range plasmids found globally in diverse species of *Gammaproteobacteria* isolated from food products, food-producing animals, and humans (1). IncC plasmids circulate not only in several pathogenic species of *Enterobacteriaceae* but also in seventh pandemic African O1 El Tor isolates of *Vibrio cholerae*, the infectious agent of the diarrheal disease cholera, and other *Vibrionaceae* (2–4). IncC plasmids may carry transposons and integrons that confer resistance to clinically important families of antibiotics, including

Citation Humbert M, Huguet KT, Coulombe F, Burrus V. 2019. Entry exclusion of conjugative plasmids of the IncA, IncC, and related untyped incompatibility groups. *J Bacteriol* 201: e00731-18. <https://doi.org/10.1128/JB.00731-18>.

Editor George O’Toole, Geisel School of Medicine at Dartmouth

Copyright © 2019 American Society for Microbiology. All Rights Reserved.

Address correspondence to Vincent Burrus, Vincent.Burrus@USherbrooke.ca.

Received 29 November 2018

Accepted 5 March 2019

Accepted manuscript posted online 11 March 2019

Published 24 April 2019

β -lactams and β -lactamase inhibitors, cephalosporins, carbapenems, aminoglycosides, tetracyclines, quinolones, trimethoprim, and sulfonamides (1, 4). IncC plasmids also participate in the emergence of drug-resistant bacteria through the mobilization of unrelated multidrug resistance-conferring genomic islands, such as MGI_{Vch}Hai6 of *V. cholerae* and *Salmonella* genomic island 1 (SGI1) (5, 6).

Plasmid incompatibility and exclusion are two distinct processes preventing a bacterium that carries a plasmid from being a suitable recipient for the acquisition of an identical or similar plasmid. Incompatibility results from the inability of similar plasmids to be propagated stably in the same cell line due to interference between their respective replication or partition functions. The process of exclusion presents a barrier against DNA transfer during conjugation between two bacterial cells carrying related elements (7). Conjugation is mediated by a type IV secretion system (T4SS), a multiprotein nanomachine that spans the cell envelope to translocate proteins and DNA from a donor cell to a recipient cell (8). Two mechanisms of the process of exclusion occurring during conjugation have been described previously (9). Surface exclusion prevents close contact between cells, whereas entry exclusion impedes DNA transfer after formation of the mating pair. Entry exclusion has been studied in F-like conjugative plasmids (10) and in plasmids of the IncHI1 (11), IncI1/IncI γ (12), IncN/IncW (13, 14), and IncP α incompatibility groups (15), as well as in integrative and conjugative elements (ICEs) of the SXT/R391 family (16). Apart from IncHI1 plasmids, these conjugative elements encode an entry exclusion (Eex) protein that is required in the recipient cell but not in the donor cell to exert entry exclusion. For F-like plasmids and SXT/R391 ICEs, the Eex protein in the recipient cell interacts with a protein of the T4SS in the donor cell, namely, the VirB6-like mating-pair stabilization protein TraG (10, 16). For IncI1/IncI γ plasmids, the Eex protein interacts with TraY in the donor cell (12). This VirB6-like protein plays the same role as TraG but was given a different name due to inconsistencies in the naming of T4SS proteins encoded by different plasmid types (17). Related conjugative elements carrying orthologous and yet divergent exclusion systems can fall into different exclusion groups that do not exclude each other (12, 18, 19). Entry exclusion has been shown to be an essential feature of conjugative plasmids and is thought to be ubiquitous among conjugative elements (20). However, the mechanisms by which entry exclusion factors interact with each other and interfere with DNA transfer are not yet understood.

IncC plasmids are closely related to IncA conjugative plasmids such as pRA1 from the fish pathogen *Aeromonas hydrophila* (21). Upon initial characterization in the early 1970s, pRA1 was found to be compatible with plasmids of all known compatibility groups, including IncC plasmids (22, 23). However, marked entry exclusion was also observed when the IncC pR57b plasmid or pIP40a plasmid was transferred to a recipient carrying the IncA plasmid pRA1 (24). At the time, these observations led Hedges (25) to combine IncA and IncC plasmids to form the "A-C complex." The term "IncA/C" was introduced later for reasons that remain unclear and has become widespread since then in the literature (1). Ambrose et al. (26) recently confirmed the compatibility of IncA and IncC replicons and recommended avoiding the use of the term "IncA/C" because it causes unnecessary confusion. Those authors also showed that the apparent "incompatibility" likely results from IncA and IncC plasmids exerting entry exclusion effects on one another, thereby suggesting the existence of a shared mechanism of exclusion. By analogy with F-like plasmids and the SXT/R391 ICEs, they proposed a small open reading frame (ORF) that is conserved in IncA and IncC plasmids and immediately downstream of *traG* as a potential candidate for an entry exclusion determinant.

Our group recently showed that SGI1 encodes three T4SS subunit proteins, TraN_S, TraH_S, and TraG_S, that are distantly related to the corresponding TraN_C, TraH_C, and TraG_C proteins encoded by IncC plasmids (27). The three protein subunits alone are insufficient to form a functional conjugative T4SS, and SGI1 requires a helper IncA or IncC plasmid for transfer (28). However, the substitution of TraG_S of SGI1 for the VirB6-like TraG_C subunits of the IncC F-type T4SS in donor cells has been shown to allow

TABLE 1 *E. coli* K-12-derivative strains, plasmids, and genomic islands used in this study

Strain, plasmid, or genomic island	Relevant genotype or phenotype ^a	Source or reference(s)
Strains		
CAG18439	MG1655 <i>lacZU118 lacI42::Tn10</i> (Tc)	54
BW25113	F [−] Δ (<i>araD-araB</i>)567 Δ (<i>lacZ4787::rrnB-3</i>) λ^- <i>rph-1</i> Δ (<i>rhaD-rhaB</i>)568 <i>hsdR514</i>	37
GG56	Nx ^r derivative of BW25113 (Nx)	55, 56
VB112	Rf ^r derivative of MG1655 (Rf)	57
Plasmids		
pAsa4c	Cm ^r conjugative plasmid from <i>A. salmonicida</i> JF2267 (Cm)	58
pVCR94	IncC conjugative plasmid from <i>V. cholerae</i> O1 El Tor (Su Tm Cm Ap Tc Sm)	59
pVCR94ΔX	Su ^r derivative of pVCR94, lacking the large fragment containing the multidrug resistance gene cluster	59
pVCR94 ^{Sp}	Sp ^r derivative of pVCR94ΔX (pVCR94ΔX2) (Sp Su)	29
pVCR94 ^{Kn}	Kn ^r derivative of pVCR94ΔX (pVCR94ΔX3) (Kn Su)	29
pVCR94 ^{Sp} ΔacaCD	pVCR94 ^{Sp} ΔacaCD (Sp Su)	29
pVCR94 ^{Sp} ΔA	pVCR94 ^{Sp} Δ(<i>vcrx004-vcrx028</i>) Δ(<i>vcrx030-vcrx059</i>) (Sp Su)	This study
pVCR94 ^{Sp} ΔB	pVCR94 ^{Sp} Δ(<i>vcrx064-vcrx069</i>) (Sp Su)	This study
pVCR94 ^{Sp} ΔC	pVCR94 ^{Sp} Δ <i>vcrx076</i> (Sp Su)	This study
pVCR94 ^{Sp} ΔD	pVCR94 ^{Sp} Δ(<i>vcrx085-vcrx141</i>) (Sp Su)	This study
pVCR94 ^{Sp} ΔE	pVCR94 ^{Sp} Δ <i>vcrx145</i> (Sp Su)	This study
pVCR94 ^{Sp} ΔtraG _C	pVCR94 ^{Sp} ΔtraG _C (Sp Su)	27
pBAD30	<i>ori_{p15A}</i> <i>araC</i> P _{BAD} (Ap)	39
peexC	pBAD30::eexC (<i>vcrx145</i> of pVCR94ΔX) (Ap)	This study
peexCrev	pBAD30::eexCrev (<i>vcrx145</i> in reverse orientation) (Ap)	This study
ptrag _S	pBAD30::traG _S (<i>traG_S</i> from SGI1) (Ap)	27
peexC2	pBAD30::eexC2 (<i>YR71pYR1_0178</i> of pYR1) (Ap)	This study
peexC3	pBAD30::eexC3 (<i>pKHM-243_0211</i> of pKHM-1) (Ap)	This study
peexC1	pBAD30::eexC1 (<i>pRA1_0150</i> of pRA1) (Ap)	This study
peexD	pBAD30::eexD (sequence 139329 to 139709 of pAhD4-1) (Ap)	This study
peexE	pBAD30::eexE (sequence 156821 to 157174 of pAsa4c) (Ap)	This study
ptrag _E	pBAD30::traG _E (sequence 153216 to 156770 of pAsa4c) (Ap)	This study
pCmE	pBAD30::eexCmE (Ap)	This study
pEmC	pBAD30::eexEmC (Ap)	This study
pClo	CloDF13::TnAΔEcoRV (pSU4628) (Ap)	60
pSIM6	Thermoinducible expression of λRed recombination (<i>ori^{ts}</i> Ap)	61
pKD3	<i>cat</i> (Cm) template for one-step chromosomal gene inactivation	37
pKD4	<i>aph</i> (Kn) template for one-step chromosomal gene inactivation	37
pTT01	<i>tetM</i> (Tc) template for one-step chromosomal gene inactivation	62
pE-FLP	Constitutive expression of FLP recombinase (<i>ori^{ts}</i> Ap)	38
Genomic islands		
SGI1	SGI1 inserted into the 3' end of <i>trmE</i> (Ap Cm Sp Sm Su Tc)	29
SGI1 ^{Kn}	ΔIn104::aph mutant of SGI1 devoid of the integron In104 (Kn)	27
SGI1 ^{Kn} ΔtraG _S	ΔIn104::aph ΔtraG _S mutant of SGI1 (Kn)	27

^aAp, ampicillin; Cm, chloramphenicol; Kn, kanamycin; Nx, nalidixic acid; Rf, rifampin; Sp, spectinomycin; Sm, streptomycin; Su, sulfamethoxazole; Tc, tetracycline; Tm, trimethoprim; ts, thermosensitive.

SGI1 transfer to a recipient strain carrying an IncC plasmid, thereby suggesting that entry exclusion cannot occur in this context (27). Since the Eex determinant of IncC and related plasmids was unknown, we investigated the exclusion system of IncC plasmids to identify the *eex* gene. Phylogenetic analyses and exclusion assays revealed that all known IncC and IncA plasmids belong to the same exclusion group and allowed us to define additional exclusion groups in more distantly related, untyped conjugative plasmids.

RESULTS

Identification of the gene mediating IncC plasmid entry exclusion in recipient cells. BLASTP analysis of the predicted translation products of all open reading frames of IncC plasmid pVCR94ΔX (GenBank accession number [KF551948.1](https://www.ncbi.nlm.nih.gov/nuclot/KF551948.1)) (Table 1) failed to reveal any significant homology with any known exclusion proteins encoded by conjugative plasmids or ICEs. To examine whether this IncC plasmid was able to mediate entry exclusion, we used two differentially marked derivatives carrying either a kanamycin resistance gene (pVCR94^{Kn}) or a spectinomycin resistance gene (pVCR94^{Sp}).

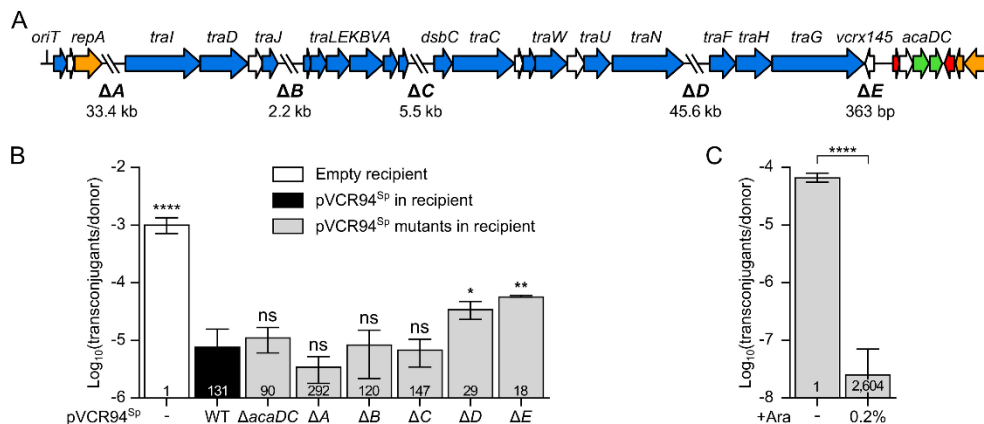


FIG 1 Mapping of IncC exclusion functions. (A) Linear schematic representation of pVCR94ΔX gene map. The positions and orientations of open reading frames (ORFs) are indicated by arrows. Colors depict the function deduced from functional analyses and BLAST comparisons as follows: orange, replication and partition; blue, conjugative transfer; red, transcriptional repression; green, transcriptional activation; white, unknown function. The positions of fragments deleted in pVCR94^{Sp} (ΔA to ΔE) are indicated by backslashes, and lengths of deletions are indicated below the schematic representation. (B) Exclusion of pClo transferred from a donor strain bearing pVCR94^{Kn} by a set of deletion mutants of pVCR94^{Sp} in recipient cells. The donor strain was *E. coli* VB112 bearing pVCR94^{Kn} and pClo (Ap^r). All recipient strains were derivatives of *E. coli* GG56 that were either plasmid free (–) or bearing pVCR94^{Sp} (wild type [WT]) or its deletion mutants. (C) Expression of *eexC* (*vcrx145*) in a recipient strain is sufficient to exclude transfer of pVCR94^{Sp}. The donor strain *E. coli* GG56 bearing pVCR94^{Sp} was mated with *E. coli* CAG18439 carrying *peexC* as the recipient without or with arabinose (+Ara). Each bar represents the mean of results from three independent experiments, with error bars indicating the standard deviations. Exclusion indices (EI) are indicated at the bottom of each bar. Statistical analyses were carried out on the logarithm of the values using one-way analysis of variance (ANOVA) with Dunnett's multiple-comparison test (B) and an unpaired *t* test (two-tailed) (C) to compare induced transfer and noninduced transfer. Statistical significance is indicated as follows: ****, $P < 0.0001$; **, $P < 0.01$; *, $P < 0.05$; ns, not significant.

pVCR94^{Kn} and pVCR94^{Sp} were introduced into a rifampin-resistant *Escherichia coli* K-12 strain and a nalidixic acid-resistant *E. coli* K-12 strain, respectively. Since incompatibility between these two virtually identical plasmids prevented the isolation of transconjugant colonies, we monitored the transfer of ampicillin-resistant, broad-host-range mobilizable plasmid pCloDF13 (pClo), which relies entirely on the mating pore of pVCR94^{Kn} for transfer (27). pClo was thus used as a proxy to assess entry exclusion occurring between the two strains, each bearing a different differentially marked plasmid. The frequency of transfer of pClo to a plasmid-free recipient strain was ~130 times higher than the frequency of transfer to an isogenic recipient strain carrying pVCR94^{Sp} (Fig. 1B). This value represents the exclusion index (EI) as defined by others (16, 20) and confirms that our pVCR94 derivatives exert entry exclusion effects on each other. In this study, we concluded that no exclusion was present when EI values were below 3.

We next used the same approach to map the pVCR94^{Sp} genes that mediate exclusion in recipient cells using recipient strains harboring a set of deletion mutations of pVCR94^{Sp}. These mutants lack either the *acaDC* transcriptional activator genes or five regions encompassing mostly genes of unknown function (ΔA to ΔE in Fig. 1A). None of the mutations was found to affect pVCR94^{Sp} stability as long as spectinomycin was used as the selective agent. Deletion of *acaDC* had no effect on exclusion (EI = 90) of pClo (Fig. 1B), suggesting that expression of entry exclusion in recipient cells is not controlled by the master activator of IncC conjugative transfer genes (29). Similar levels of exclusion were also observed using the ΔA, ΔB, and ΔC mutants (EI > 120), showing that none of those three regions contains exclusion genes. Although the frequency of transfer of pClo was not restored to the level observed in using a plasmid-free recipient strain, exclusion was reduced when the recipient carried either the ΔD (EI = 29) or ΔE (EI = 18) mutants of pVCR94^{Sp} (Fig. 1B). This result indicates that the absence of region D or region E in the recipient facilitates the transfer of pClo, either by directly abolishing an exclusion mechanism or by destabilizing pVCR94^{Sp} in the recipient cells, thereby weakening exclusion.

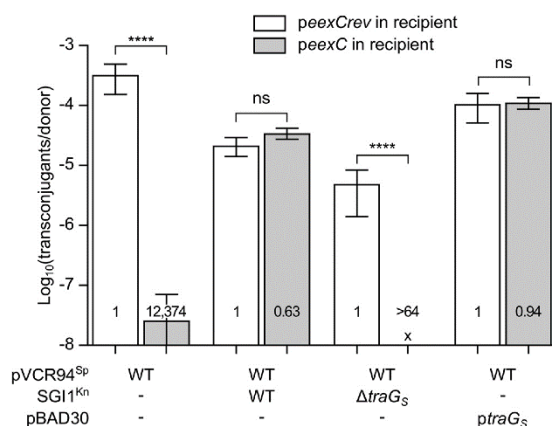


FIG 2 TraG_S of SGI1 abolishes EexC-mediated IncC entry exclusion. Mating assays were carried out using as donors GG56 bearing pVCR94^{Sp} alone or in association with SGI1^{Kn} or its ΔtraG_S mutant or ptraG_S. CAG18439 strains bearing peexC or peexCrev (negative control that contains eexC in the reverse orientation relative to the P_{BAD} promoter) were used as recipient strains. Each bar represents the mean of results from three independent experiments, with error bars indicating standard deviations. Exclusion indices (EI) are indicated at the bottom of each bar. Statistical analyses were carried out on the logarithm of the values using one-way ANOVA with Sidak's posttest to compare each bar to its corresponding control. "x" indicates that the frequency of transfer was below the detection limit (<10⁻⁸). ****, *P* < 0.0001; ns, not significant.

Since a single open reading frame, *vcrx145*, is present in the deletion region Δ*E*, further experiments were carried out to determine its function. By analogy with SXT/R391 ICEs and F-like plasmids, this small ORF downstream of *traG* (Fig. 1A) was recently proposed as a potential candidate for an entry exclusion determinant in IncC and IncA plasmids (26). *vcrx145* was cloned under the control of the arabinose-inducible P_{BAD} promoter and introduced into a tetracycline-resistant (Tc^r) *E. coli* recipient strain to test whether its expression was sufficient to mediate exclusion. The frequency of transfer of pVCR94^{Sp} to the recipient strain expressing *vcrx145* was considerably lower (EI = 2,604) than the frequency of transfer to the same recipient strain in the absence of arabinose (Fig. 1C). Therefore, expression of *vcrx145* in recipient cells is sufficient to mediate entry exclusion of pVCR94^{Sp}. On the basis of these observations and evidence presented below, we renamed *vcrx145* "eexC" (for "entry exclusion of IncC plasmids").

TraG is the entry exclusion factor in donor cells. We have previously reported that SGI1 encodes three T4SS subunits, namely, TraN_S, TraH_S and TraG_S, that are expressed under the control of the IncC-encoded AcaCD master activator (27). We found that despite weak relatedness (37% identity), the TraG_S subunit encoded by SGI1 displaces the IncC-encoded TraG_C in the T4SS of IncC plasmids (27). Consequently, SGI1 transfers efficiently between donor and recipient cells that both carry an IncC plasmid. This observation suggests that TraG_C in the donor cells and EexC in the recipient cells are interacting partners that together mediate entry exclusion. To test this hypothesis, we monitored the transfer of pVCR94^{Sp} from donor strains carrying or devoid of a chromosomal copy of SGI1^{Kn} into a recipient strain expressing eexC. SGI1^{Kn} is a kanamycin-resistant derivative of SGI1 that lacks the multidrug resistance complex integron In104 (27).

When SGI1^{Kn} was present in the donor strain, exclusion was deficient (EI = 0.63), as pVCR94^{Sp} transfer rates were virtually identical between the recipients expressing eexC and those not expressing eexC (Fig. 2). In contrast, when SGI1^{Kn}ΔtraG_S was used in the donor, transfer of pVCR94^{Sp} was excluded by recipients expressing eexC (EI > 64). As reported previously (27), transfer of pVCR94^{Sp} was also impaired by the coresident SGI1^{Kn} (150-fold reduction compared to a donor lacking SGI1^{Kn}). This impairment is

likely due to the alteration of the T4SS by the SGI1-encoded TraH_S subunit that displaces the pVCR94^{Sp}-encoded TraH_C subunit (64% identity between TraH subunits) (27). TraH subunits are localized in the periplasm/outer membrane and predicted to be involved in mating apparatus assembly and stabilization (30). Therefore, to confirm that TraG_S alone could alleviate exclusion mediated by EexC interacting with TraG_C, we used a donor strain devoid of SGI1^{Kn} and containing *ptrag_S* to express only *traG_S*. When TraG_S was produced in the donor, no exclusion was exerted toward the IncC plasmid (EI = 0.94) (Fig. 2). Taken together, these results strongly suggest that TraG_C interacts with EexC to mediate exclusion. Considering that, like other F-type Eex/TraG pairs, EexC and TraG_C are predicted to be inner membrane proteins (PSORT-B cytoplasmic membrane scores of 10 and 4.12, respectively) and that TraG_C is required in the donor while EexC acts in the recipient, the exclusion mechanism is most likely entry exclusion rather than surface exclusion (10, 16, 20).

Clustering of EexC and TraG_C orthologues suggests multiple-entry exclusion groups. Conjugative elements of the same family can belong to different entry exclusion groups. For instance, F-like plasmids F and R100 do not exclude each other (18). Likewise, cells bearing the ICE SXT inhibit the entry of a second copy of the ICE, but not R391, and vice versa (16). To assess the diversity of entry exclusion groups and determine their specificity among genetic elements related to IncC conjugative plasmids, we searched the GenBank database for EexC and TraG_C homologs. In several instances, such as pAhD4-1 of *Aeromonas hydrophila*, *eex* loci were not annotated in the nucleotide sequences. Therefore, we performed searches using tblastn (see Data Sets S1 and S2 in the supplemental material). A phylogenetic analysis of the unique representatives of EexC and TraG_C homologs was then carried out. In both the TraG and Eex trees, IncA and IncC plasmids clustered together (Fig. 3A and B), which is consistent with recent data confirming that IncC and IncA plasmids exclude each other and therefore belong to the same exclusion group (26). In fact, Eex homologs encoded by all IncC and IncA plasmids seem to form a unique lineage of EexC-like proteins that also contain putative Eex proteins encoded by untyped plasmids pHDAS3.2 and pHDZA1.1 of the aphid endosymbiont "*Candidatus* Hamiltonella defensa" and pKMH-1 of *Citrobacter freundii* (Fig. 3A). This analysis also revealed that plasmids such as pVCR94, pMR0211, pKP-Gr642, and pEI-MS-17-156-1, which code for strictly identical TraG_C proteins (Fig. 3B; see also Data Set S1), can encode relatively divergent EexC proteins (Fig. 3A), suggesting that the small *eex* genes are evolving at a higher rate than their cognate *traG* partners. However, *eexC* genes exhibit very little diversity among IncC plasmids, as more than 96% of these plasmids in our data set (254 of 263) encode the same EexC protein as pVCR94 (Data Set S2). A strictly identical EexC protein is also encoded by non-IncC elements, such as the hybrid IncC/IncX3 plasmid pSL131_IncA/C-IncX3 of *Salmonella enterica* subsp. *enterica* serotype Lomita and the IncFIB plasmid pYT3 of *S. Typhimurium*, or by those present on the chromosome of *Providencia stuartii*, *S. Newport*, and *S. Typhimurium* (Data Set S2). The 8 IncA plasmids included in our data set encode either the EexC1 variant or the EexC4 variant, which seem to be specific to IncA plasmids, as neither variant was found to be encoded by any IncC plasmids (Data Set S2). Nevertheless, we predict that all the conjugative elements cited above belong to entry exclusion group C (EexC/TraG_C).

Other untyped plasmids, including pAsa4-like or pAhD4-1-like plasmids found in *Aeromonas* species, encode more distantly related Eex and TraG proteins that form distinct lineages (Fig. 3A and B). However, the strong divergence of these pairs of proteins suggests that two distinct entry exclusion groups exist within this lineage: pAsa4c-like plasmids likely form exclusion group E (EexE/TraG_E) whereas pAhD4-1-like plasmids would form exclusion group D (EexD/TraG_D).

Confirmation of exclusion groups by mating assays. Since three different pairs of Eex/TraG exclusion partners were identified, we hypothesized that EexC proteins would exclude pVCR94^{Sp} transfer and not pAsa4c transfer and that EexE would exclude pAsa4c transfer and not pVCR94^{Sp} transfer. We also hypothesized that EexD should

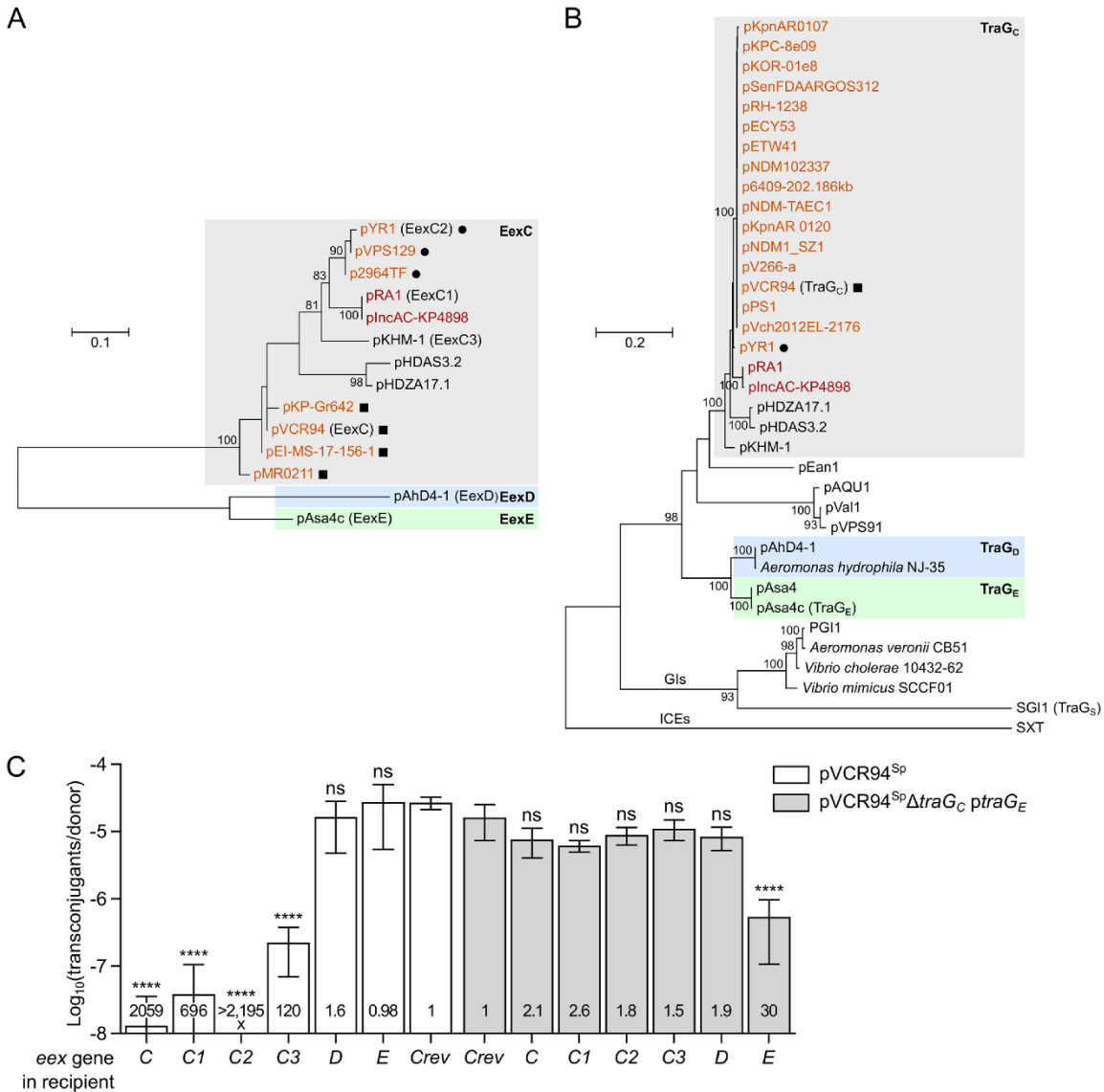


FIG 3 Multiple-entry exclusion groups among IncC-related conjugative plasmids. (A and B) Maximum likelihood phylogenetic analysis of Eex (A) and TraG (B) homologs. Trees with the highest log likelihoods (−987.74 and −16,264.82 for Eex and TraG, respectively) are shown. Bootstrap supports are indicated as percentages at the branching points only when >80%. Branch length represents the number of substitutions per site over 114 and 1,204 amino acid positions for Eex and TraG proteins, respectively. Only one representative per cluster of identical proteins is shown in each tree (Data Sets S1 and S2). IncC and IncA plasmids are shown in orange and red, respectively. Circles and squares indicate plasmids coding for identical TraG proteins. GlS indicates a lineage of TraG proteins encoded by SGI1-like genomic islands. TraG of the ICE SXT (TraG/EexS exclusion system) was used as the outgroup. (C) Experimental confirmation of exclusion groups. Mating assays were performed using as donor strains GG56 bearing either pVCR94^{Sp} or its ΔtraG_C mutant with ptrag_E. Recipient strains were CAG18439 strains expressing the different putative eex genes or carrying pexCrev used as the nonexclusion control. Each bar represents the mean of results from three independent experiments, with error bars indicating the standard deviations. Exclusion indexes (EI) are indicated at the bottom of each bar. One-way ANOVAs with Dunnett's multiple-comparison test were carried out on the logarithm of the values to compare each bar to the nonexclusion control for each donor. ****, $P < 0.0001$; ns, not significant.

exclude neither. To confirm the existence of three distinct exclusion groups, we cloned five additional eex genes originating from IncA pRA1 (eexC1), IncC pYR1 (eexC2), and untyped plasmids pKHM-1 (eexC3), pAsa4c (eexE) and pAhD4-1 (eexD), under the control of the P_{BAD} promoter and then assessed whether their expression in a recipient

strain excluded the transfer of pVCR94^{Sp} or pAsa4c. Expression of any of the four *eexC* alleles (including the IncA-derived *eexC1*) strongly inhibited or even abolished transfer of pVCR94^{Sp} (EI > 120), confirming that IncC and IncA plasmids belong to exclusion group C (Fig. 3C). In contrast, EexE and EexD were unable to exclude pVCR94^{Sp}, as shown by the low EIs that were not statistically different from the control, confirming that neither protein belongs to exclusion group C.

Direct assessment of pAsa4c exclusion was challenging using *E. coli* as a host. Despite multiple attempts, intraspecific *E. coli* transfer of pAsa4c remained near the detection threshold (10^{-8} transconjugant/donor), even using different donor and recipient strains and temperatures closer to the optimal growth temperature of its original host, the fish pathogen *Aeromonas salmonicida* subsp. *salmonicida*. This observation suggests that physiological differences limit the transfer of pAsa4c in *E. coli*. Therefore, we used instead a donor strain bearing pVCR94^{Sp} Δ traG_C complemented with traG_E of pAsa4c expressed from ptrG_E and then monitored the transfer into the recipient strains expressing the different *eex* genes. Using this traG substitution, we observed that transfer of pVCR94^{Sp} and its Δ traG_C mutant complemented with traG_E occurred at frequencies that were indistinguishable (Fig. 3C, compare *Crev* bars). Moreover, these assays revealed that exclusion of the complemented mutant occurred only when *eexE* was expressed in the recipient (EI = 30) (Fig. 3C). None of the four *eexC* variants or *eexD* was able to exclude pVCR94^{Sp} Δ traG_C complemented with traG_E (EI ranging between 1.5 and 2.6). This result confirms that EexC and TraG_C of the IncA and IncC plasmids form an entry exclusion group distinct from the EexE/TraG_E system of untyped pAsa4c. Furthermore, since neither TraG_C nor TraG_E showed an interaction with EexD from pAHD4-1 (EI = 1.6 or 1.9, respectively), this plasmid likely belongs to a third distinct entry exclusion group.

The variable C terminus of EexC defines the specificity of entry exclusion. BLASTP alignment of EexC and EexE revealed that the two proteins share only 52% identity over 83% coverage. However, the 62 amino acid (aa) residues of the N terminus of both proteins share 69% identity. The predicted secondary structures of both N termini are strikingly similar and consist of an α -helix folding that is predicted to be located in the inner membrane (PSORTb cytoplasmic membrane scores of 10.0 and 9.82 for EexC and EexE, respectively) (Fig. 4A). Comparison of the hydropathy profiles of the two proteins also shows strikingly similar features, particularly over the 80 N-terminal amino acid residues (Fig. 4B). These observations suggest that specificity of entry exclusion is defined by the highly variable C terminus of both Eex proteins. To test this hypothesis, we swapped the C termini of EexC at position 82 and EexE at position 79 (Fig. 4A) and then assessed whether the expression of the two resulting chimeric genes, *EmC* and *CmE*, in a recipient strain would exclude pVCR94^{Sp} or the corresponding Δ traG_C mutant complemented with traG_E. We observed that *CmE*, which contains the C terminus of EexE, did not exclude pVCR94^{Sp} (EI = 0.19) whereas it excluded pVCR94^{Sp} Δ traG_C complemented with traG_E (EI = 37) (Fig. 4C). Conversely, *EmC*, which contains the C terminus of EexC, excluded pVCR94^{Sp} (EI = 49) but not pVCR94^{Sp} Δ traG_C complemented with traG_E (EI = 0.5). Furthermore, exclusion activities mediated by the C terminus of the chimeras toward the cognate TraG were comparable to those seen with the native proteins (Fig. 4C; compare C to *EmC* and E to *CmE*). Together, these results confirm that the C termini of EexE and EexC define entry exclusion specificity.

DISCUSSION

Entry exclusion is a feature conserved among conjugative plasmids of different incompatibility groups and has also been reported in one family of ICEs to date (20). Entry exclusion has been proposed to confer immunity to lethal zygotosis, i.e., death of recipient cells resulting from exposure to an excess of donor cells that could increase cell permeability by inducing extensive damage in the recipient membrane. It also likely increases donor fitness by limiting excessive energy consumption due to transfer to another donor. Moreover, entry exclusion improves plasmid stability by impeding entry of highly similar, incompatible replicons that can lead to plasmid displacement and

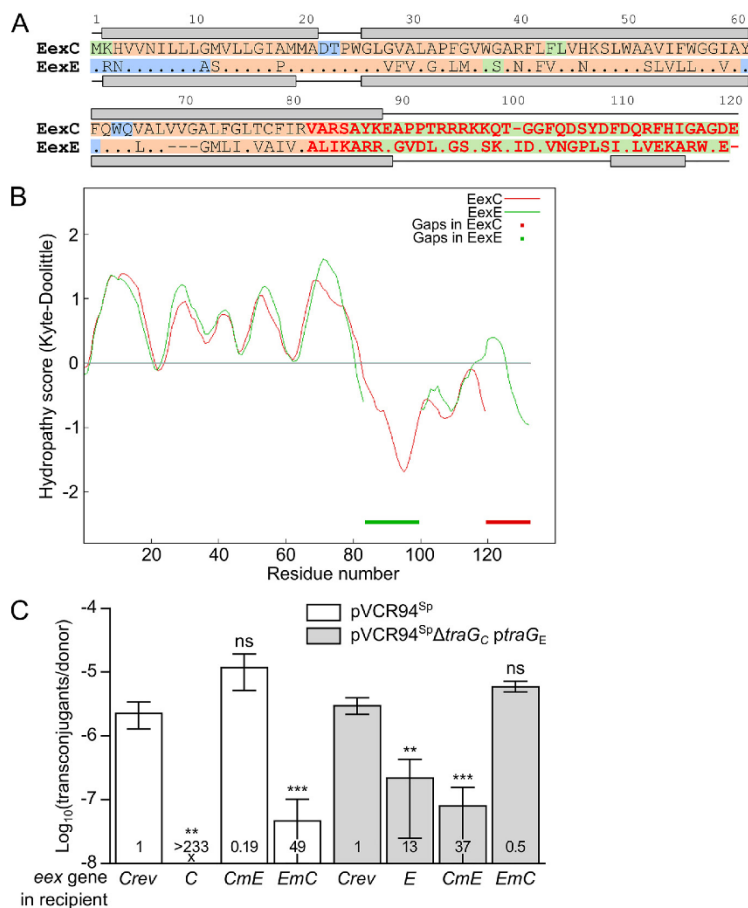


FIG 4 Localization of the specificity domain of Eex proteins. (A) MUSCLE alignment of EexC and EexE. EexC was used as the reference, and identical amino acids in EexE are represented by a dot. Predicted secondary structures are shown above and below the sequence of each protein, with lines and gray boxes representing coils and helices, respectively. Background colors indicate amino acid residues predicted to be cytoplasmic (green) or periplasmic (blue) or present within a transmembrane helix (orange). Variable C termini of EexC and EexE shown in red were swapped to construct the chimeric proteins CmE and EmC. (B) Aligned Kyte-Doolittle hydropathy plots of EexC and EexE. Thick lines indicate alignment gaps in EexC and EexE. (C) Specificity of Eex chimeras toward TraG_C and TraG_E. Mating assays were performed using as the donor strains GG56 bearing either pVCR94^{Sp} or its ΔtraG_C mutant with ptraG_E. All recipient strains were derivatives of CAG18439 expressing eexC or eexE or either of the two chimeras eexCmE and eexEmC. CAG18439 bearing pVCR94^{Sp} was used as the nonexclusion control. Each bar represents the mean of results from three independent experiments, with error bars indicating the standard deviations. Exclusion indexes (EI) are indicated at the bottom of each bar. One-way ANOVAs with Dunnett's multiple-comparison test were carried out on the logarithm of the values to compare each bar to the nonexclusion control for each donor. ***, $P < 0.001$; **, $P < 0.01$; ns, not significant.

loss. IncC and IncA plasmids were long suspected to exclude each other's entry, but the genetic determinants responsible for exclusion were unknown (22–25). Confusion between exclusion and incompatibility has plagued the nomenclature and typing of these plasmids since the late 1980s (26, 31). Nevertheless, Ambrose et al. (26) recently confirmed that IncA and IncC replicons are compatible and that the apparent incompatibility between these two plasmid groups most likely resulted from entry exclusion. Here, we identified the genes involved in entry exclusion of IncA and IncC plasmids and showed that the product of eexC in recipient cells recognizes the cognate TraG_C protein in donor cells, exerting entry exclusion toward both IncC and IncA plasmids. Therefore,

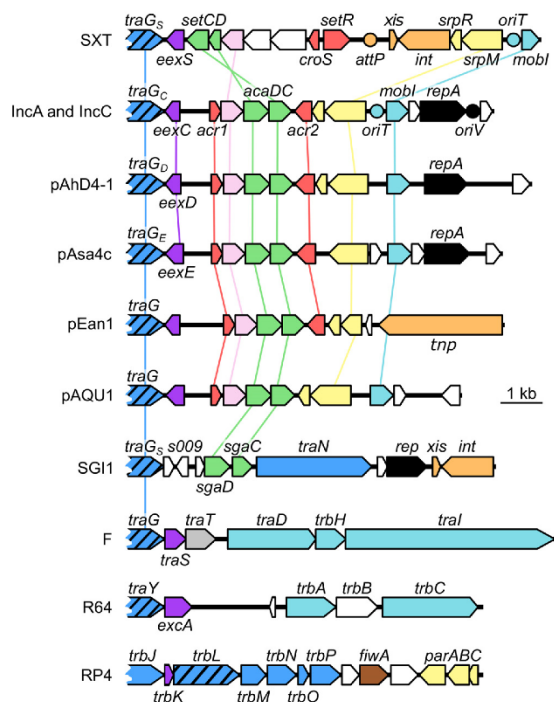


FIG 5 Comparison of the genetic contexts of entry exclusion genes in diverse mobile genetic elements. A schematic representation of demonstrated or putative exclusion loci of the ICE SXT, IncC plasmid pVCR94; untyped plasmids pAhD4-1, pAsa4c, pEan1 and pAQU1; genomic island SGI1; IncF1 plasmid F; IncI1 plasmid R64; and IncP α plasmid RP4 is shown. Arrows of identical colors represent genes predicted to have similar functions and are color coded as follows: dark blue, T4SS subunit; hatched blue, VirB6-like subunit; light blue, DNA processing; violet, entry exclusion; gray, surface exclusion; green, transcriptional activator; pink, lytic transglycosylase; red, transcriptional repressor; yellow, partition; orange, site-specific recombination and transposition; black, replication; white, unknown function. Circles indicate the position of origins of transfer (*oriT*), of the origin of replication (*oriV*), of pVCR94 based on identity with pRA1, and of the *attP* site used for chromosomal integration of SXT by site-specific recombination.

the C entry exclusion group spans at least two different plasmid incompatibility groups, a discrepancy that led to the grouping of IncA and IncC plasmids into the incorrect and now obsolete "IncA/C" group, alias the "A/C" group (1, 26). Moreover, at least one IncFIB plasmid, pYT3, also seems to belong to the C exclusion group (see Data Sets S1 and S2 in the supplemental material).

The identification of EexC revealed the existence of homologous exclusion proteins, such as EexE and EexD, that are encoded by untyped conjugative plasmids that are related either closely (e.g., pAsa4c) or more distantly (e.g., pAhD4-1) to IncA and IncC plasmids. These Eex proteins define additional exclusion groups that do not exclude IncC and IncA plasmids. Furthermore, several genomic islands and plasmids such as pEan1 or pVPS91 were found to code for closely related TraG homologs and yet encode predicted Eex proteins related distantly (<40% identity) to EexC, EexD, or EexE (Fig. 3A and B; see also Fig. 5). Consistent with this observation, EexC has no detectable sequence similarity to EexS of SXT, TraS of the F-like plasmids F and R100, or ExcA of the IncI1/IncIy plasmids R64 and R621a (Fig. 5). The lack of homology between Eex proteins from one entry exclusion system and those from another suggests that *eex* genes evolve faster than most *tra* genes, including *virB6* homologs such as *traG* (20). According to Sakuma et al. (12), the association of VirB6-like proteins within a high-order multimeric structure that requires interactions with several other transfer proteins in the T4SS imposes a higher level of conservation than Eex proteins, which likely

interact only with their cognate TraG partner. While SGI1 TraG_S is a functional substitute for IncC TraG_C despite being only 37% identical, no entry exclusion gene has been identified in SGI1 to date. Nevertheless, we predict that SGI1 *s009* codes for an Eex protein on the basis of its location relative to *traG_S* and of the partial identity of its translation product with F TraS (38% identity and 69% similarity of aa 45 to 57 of S009 to aa 50 to 62 of TraS) (Fig. 5).

Entry exclusion systems of IncC-like, F-like, and Inc11/Inc1γ plasmids as well as SXT/R391 ICEs exhibit striking functional similarities. In all these systems, exclusion involves the VirB6-like protein TraG (or TraY for Inc11/Inc1γ plasmids) in the donor and Eex in the recipient (12, 16, 18). F and SXT TraG proteins have been shown to be polytopic inner membrane proteins, with large hydrophilic C-terminal domains that are predicted to localize in the cytoplasm (10, 16). Like TraY (12), TraG_C is predicted to be an inner membrane protein. Though phylogenetically unrelated, F TraS, R64 ExcA, and SXT EexS have all been shown or predicted to be inner membrane proteins (10, 12, 16). Like their counterparts, EexE, EexD, and EexC are predicted to be bound to the inner membrane. The segments of TraG_S and EexS of SXT that mediate entry exclusion specificity were shown to be cytoplasmic, thereby suggesting a model of entry exclusion in which a cytoplasmic segment of TraG would be translocated into the recipient cell cytoplasm (32). Indeed, exclusion specificity was shown to be controlled by the C-terminal parts of SXT EexS and of closely related R391 EexR and only three amino acid residues (aa 606 to 608) located in the cytoplasmic, central region of their cognate TraG (16, 32). Precise mapping of the amino acid residues that control exclusion specificity was facilitated by the low divergence between TraG_S and TraG_R (98% identity) (16). The TraS exclusion proteins of the F-like plasmids F and R100 share only 17% identity and lack regions of higher dissimilarity to aid in predicting the features that are recognized by TraG (10, 18). Moreover, attempts to construct stable TraS chimeras failed, preventing the identification of the region of TraS responsible for exclusion specificity. Audette et al. (10) were, however, able to locate the segment of TraG that interacts with TraS within a discrete segment (aa 610 to 673) on the basis of its lower conservation (55.7% identity versus 93% overall identity) and the use of TraG chimera proteins. For IncC-like plasmids, TraG_C and TraG_E share only 76% overall identity, with 89% identity over the predicted inner membrane segment (aa 1 to 456) and only 68% identity over the predicted cytoplasmic segment (aa 457 to 752). Poor conservation of TraG cytoplasmic segments could hamper the mapping of the amino acid residues interacting with Eex, especially if different segments of TraG_C and TraG_E are recognized by their cognate EexC and EexE proteins. Such a feature has been reported for the entry exclusion system of Inc11/Inc1γ plasmids. Indeed, the internal variable segment of R64 TraY has been shown to recognize the R64 ExcA whereas the C-terminal variable segment of R621a TraY recognizes R621a ExcA (12). Here, we mapped the region that controls exclusion specificity to the C-terminal third of EexC and EexE. Swapping of shorter fragments could help identify the exact amino acid residues responsible for entry exclusion specificity. However, the strong divergence between the C-terminal regions of EexC and EexE (Fig. 4A) suggests that these amino acid residues might not be equivalent or even contiguous in both proteins, thereby precluding precise identification by fragment swapping. Alternative methods such as cross-linking or immunoprecipitation could be attempted, although such approaches have not been successful in attempts to decipher the interactions between TraS and TraG of F and R100 (10). Comparisons of structural data and three-dimensional (3D) models of EexC and EexE, which are not currently available, would help shed new light on the specific interactions between Eex proteins and their cognate TraG partners and provide a better understanding of the mechanism of entry exclusion, which remains to be established.

We found an unexpected lack of diversity of *eexC* variants in IncC plasmids, considering the diversity of geographical locations and bacterial species from which the plasmids of our data set originated. IncC plasmids have been typed in two main groups, type 1 and type 2, based on differences within the resistance gene clusters or loss of backbone regions adjacent to them (1). Hybrid plasmids that share features of both

types have also been described, as well as subtypes 1a and 1b, which differ by the presence or absence of a 14.5-kb diverged segment (31). We found no clear correlation between IncC plasmid type and the variant of the *eexC* gene that they carry. In fact, plasmids of all IncC types and subtypes most frequently carry the variant *eexC* (Data Set S2). Among 280 IncC plasmids, only 12 distinct EexC variants were found. In striking contrast, SXT/R391 ICEs encode a much more diverse set of Eex proteins, with two exclusion groups, S and R (16, 33). Among 31 ICEs of the S group, 18 distinct EexS variants have been described to date. Among those ICEs, 32% encode the same EexS2 protein (19, 33). A total of 19 distinct EexR variants have also been described among 31 ICEs of the R group, and the EexR1 protein variant is encoded by a small majority of only 19% of these ICEs. One hypothesis that could explain the different distribution patterns of *eexC* and *eexS/R* variants proposes a relatively clonal expansion of drug resistance-associated IncC plasmids in pathogenic species. In contrast, acquisition of SXT/R391 ICEs seems to have occurred multiple times in recent history from diverse reservoirs, especially in the seventh-pandemic *V. cholerae* lineage (34).

In our setup using pVCR94 derivatives, we observed that overexpression of different *eexC* alleles in recipients yielded different levels of exclusion effects on *traG_C*, ranging from complete reduction (*eexC2*) to only 120-fold reduction (*eexC3*) (Fig. 3C). This suggests that not all pairs of *eexC* and *traG_C* alleles confer the same levels of exclusion. A similar observation was reported for the ICE R391, which was strongly excluded by a recipient expressing the variant EexR1 and weakly excluded when variant EexR or variant EexR3 was used (19). Assessment of the exclusion strength of the 12 distinct *eexC* variants identified here against each of the *traG_C* variants identified in IncC plasmids could provide valuable information on the interactions between these two proteins and the amino acid residues involved in recognition and the specificity of exclusion. pVCR94^Δ*traG_C*, which can be complemented by *traG* genes as distant as *traG_E* or *traG_S*, could be an interesting platform for such a study to prevent inconsistency due to the variable segments present in different IncC plasmids.

In experiments reported by Ambrose et al. (26), assays of mating between strains bearing the IncA plasmid pRA1 (*eexC1/traG_{C1}*) and the IncC plasmid pRMH760 (*eexC/traG_C*) did not yield any transconjugants, thereby suggesting complete exclusion. In contrast, Datta and Hedges (24) reported that pRA1 in a recipient excluded transfer of pIP40a (*eexC/traG_C*) by only about 10-fold. Clearly, different plasmids carrying identical *eex/traG* pairs do not exert similar levels of exclusion. This observation suggests that other IncC-encoded factors may influence exclusion. This hypothesis is supported by results from our exclusion assay performed using the ΔD mutant of pVCR94^Δ. We found that pClo mobilization was more efficient toward a recipient bearing pVCR94^Δ ΔD or pVCR94^Δ ΔE ($\Delta eexC$) but that neither deletion was sufficient to restore the observed rate of transfer toward an empty recipient (Fig. 1B), thereby suggesting that the *eexC/traG_C* system is not the only exclusion system encoded by IncC plasmids. One or several of the 57 genes that contained the missing ΔD segment seem to reduce the transfer of pClo when present in the recipient strain. This reduction could have been the result of surface exclusion, like that mediated by TraT of plasmid F (35). However, since IncC plasmids do not seem to encode any TraT homolog, further investigations of the phenotype resulting from the ΔD mutation will be necessary to decipher its mechanism and how this observation relates to data published by others (24, 26).

MATERIALS AND METHODS

Bacterial strains and media. Bacterial strains and plasmids used in this study are described in Table 1. Strains were routinely grown in lysogeny broth (LB-Miller; EMD) at 37°C in an orbital shaker/incubator and were preserved at –80°C in LB broth containing 15% (vol/vol) glycerol. The following antibiotics and concentrations were used: ampicillin (Ap), 100 μ g/ml; chloramphenicol (Cm), 20 μ g/ml; kanamycin (Kn), 50 μ g/ml; nalidixic acid (Nx), 40 μ g/ml; spectinomycin (Sp), 50 μ g/ml; tetracycline (Tc), 12 μ g/ml; rifampin (Rf), 50 μ g/ml.

Mating assays. Conjugation assays were performed by mixing 100 μ l of donor cells and 100 μ l of recipient cells (typically $\sim 2 \times 10^9$ cells/ml each) that were grown overnight in LB broth at 37°C with suitable antibiotics to ensure retention of plasmids and/or SGI1 derivatives. Cells were pelleted by centrifugation for 3 min at $1,200 \times g$, washed once in 200 μ l of LB broth, and resuspended in 10 μ l of LB

TABLE 2 DNA sequences of the primers used in this study

Primer name	Nucleotide sequence (5' to 3') ^a
vcx076del.f	CTTATTTGCTCAAAAAGGGCACTTCCACAGCCCATGCGGAATAGGAACCTCAAGAT
vcx076del.r	TCTGAACCGGCTTTGAGAGGAATGAACCTTAGTTCAGGAACCTCAGAGCGCTT
94del145.for	ATGCCGTGGTAACTGGAGGTAAAGAGTAAGGGGGTTGATgtgtaggtggagctgttc
94del145.rev	ACCAACCAAATAATAAGGGGGCCAGCAGGCCCTTATTACATATGAATATCCTCCTTA
94vcx145.for	NNNNNNGAATTC AAGGAGGAATAATAATGAAACATGTGGTCAATATTTCT
94vcx145.rev	NNNNNNGAATTC TATTTCGTCTCCAGTCCAA
pAsa4cTrxA.f	GTACGAATTC AAGGAGGAATAATAATTGAGGACCTGCCTGTAT
pAsa4cTrxA.r	AGGTGAATTC TATTCACCCCAACGAGCT
ptrxAmut.f	ATGAGGAATGTGGTTAAC
ptrxAmut.r	TTATTATTCCTCCTTGAATTCG
pAsa4cTraG.f	NNNNNNGAATTC AAGGAGGAATAATAATGGGTACGTTCAACATCTA
pAsa4cTraG.r	NNNNNNGAATTC TTAATGCTTCAATCTAAGCTT
eexC183.f	TGGTTGCCATTGTCTGTCGCGCGCTCAGCATATAAAGAG
eexC279.r	CTCTTTATATGCTGACCGCGCGGCACGAACATGGCAACCA
eexC280.f	CTAACGTGCTTTATCCGCTTCTCATAAAGCGGAGAGAG
eexC182.r	CTCTCTCTCGCCTTATGAGAACGCGGATAAAGCACGTTAG
ptrxA2.f	GTCAGAATTC AAGAGGAGAGCGTGATGAG
toxdel-F	GGTCTGCAATTGGGATCATTACTTCATAAAACCTTTCTCTGTAGGTGGAGTGCTTC
toxdel-R	CAACATGAAGAGCTTCGGGTCCAAAGCAAAGGAGATGCCATGGTCCATATGAATATCCT
Del1-Kan-F	CTGCACGAACACCATCTGATATGTGGGAACGCTGCACGAAGATCCCTCAGCTGCCG
Del1-Kan-R	GTGCTCTTAACGCCCCCGAGCAAAAGCCTCAGCCCCAGATCAGAAGAACTCGTCAAGAA
Alt-del2-F	TGCATCTGTGCGACATCCATTGAATGTCTACAGGTGATAAGATCCCTCAGCTGCCG
new-Alt-del2-R	CCGAATGTTTTTCGAAGCGCAACGCTAATCTTATTGGCTCAGAAGAACTCGTCAAGAA
new-del4-F	ACTCTGCTTTTCCCTGATCAAGAAGGTAAGGGAAAGCAACGTCGTGACTGGGAAACCT
new-del4-R	CTTCTGCGCTTCTGCGCCATCTTCATCTGCTTCTGACACTCCAGCATATTCAGCGAC

^aEcoRI restriction sites are underlined.

broth. Mating mixtures were then deposited as drops on LB agar plates and incubated at 37°C for 2 h. The cells were recovered from the plates in 800 μ l of LB broth, subjected to vortex mixing, and diluted via serial 10-fold dilutions before plating on LB agar containing suitable antibiotics was performed. Donor, recipient, and transconjugant colonies were selected using antibiotics as described in corresponding figure legends. To induce the expression of *traG* and *eex* genes, as well as *eex* chimeric genes, in complementation or expression assays, mating was carried out on LB agar plates with 0.2% arabinose. Frequencies of transfer were expressed as the number of transconjugant CFU per donor CFU from data obtained from three independent mating experiments. Exclusion indexes (EI) were calculated as the ratio of an element's frequency of transfer to a plasmid-free recipient to its frequency of transfer to the tested recipient, as defined elsewhere (16, 20).

Molecular biology methods. Plasmid DNA was extracted using an EZ-10 spin column plasmid DNA miniprep kit (Bio Basic) according to the manufacturer's instructions. All enzymes used in this study were purchased from New England Biolabs. PCR assays were performed with the primers described in Table 2. PCR conditions were as follows: (i) 3 min at 94°C for initial denaturation; (ii) 30 cycles of 10 s at 98°C for denaturation, 30 s at the appropriate annealing temperature, and 30 s/kb at 72°C for elongation; and (iii) 5 min at 72°C for the final elongation step. When necessary, PCR products were purified using an EZ-10 spin column PCR products purification kit (Bio Basic) according to the manufacturer's instructions. *E. coli* was transformed by electroporation as described by Dower et al. (36) in a Bio-Rad GenePulser Xcell apparatus set at 25 μ F, 200 V, and 1.8 kV using 1-mm-gap electroporation cuvettes.

Plasmid and strain construction. Strains, plasmids, and genomic islands used in this study are listed in Table 1. Primers are listed in Table 2.

Deletions ΔA to ΔE in pVCR94^{SP} were constructed using the one-step chromosomal gene inactivation technique with pSIM6 (37). ΔC and ΔE deletions were obtained using primer pair vcx076del.f/vcx076del.r and primer pair 94del145.for/94del145.rev, respectively, and pKD4 or pKD3 as the templates. ΔA , ΔB , and ΔD deletions were constructed in two steps. First, the *vcx028* toxin gene was deleted using primer pair toxdel-F/toxdel-R and pKD4 as the template. The FLP recombination target (FRT)-flanked kanamycin resistance cassette was then removed by Flp-catalyzed excision using pE-FLP (38). Then, deletions ΔA , ΔB , and ΔD were obtained using primer pair Del1-Kan-F/Del1-Kan-R, primer pair Alt-del2-F/new-Alt-del2-R, and primer pair new-del4-F/new-del4-R, respectively, and pKD4 (deletions ΔA and ΔB) or pTT01 (deletion ΔD) as the template. All deletions were verified by PCR and antibiotic resistance profiling.

Plasmids *peexC* and *peexCrev* were constructed by PCR amplification of *eexC* using EcoRI-containing primer pair 94vcx145.for/94vcx145.rev and pVCR94^{SP} as the template followed by EcoRI digestion of the amplified fragment and then cloning into EcoRI-digested pBAD30. Forward (*peexC*) and reverse (*peexCrev*) orientations of *eexC* relative to the *P*_{BAD} promoter were verified by *AfeI*/*AflII* digestion. Plasmids *peexE* and *ptrA_E* were constructed by PCR amplification of *eexE* and *traG_E* using pAsa4c as the template and EcoRI-containing primer pair pAsa4cTrxA.f/pAsa4cTrxA.r and primer pair pAsa4cTraG.f/pAsa4cTraG.r, respectively. The amplified fragments were then digested by EcoRI and cloned into EcoRI-digested pBAD30. *peexE* was subjected to site-directed mutagenesis to remove a 78-bp untranslated region with

primer pair ptrxAmut.f/ptrxAmut.r using Q5 site-directed mutagenesis (New England BioLabs) according to the manufacturer's instructions.

eexD, *eexC1*, *eexC2* and *eexC3* were obtained by DNA synthesis (Bio Basic) of the respective open reading frames from pAhD4-1 (CP013966.1), pRA1 (FJ705807.1), pYR1 (CP000602.1), and pKHM-1 (AP014939.1) preceded by a Shine-Dalgarno sequence (5'-AAGGAGGAATAA-3') and flanked by EcoRI and SalI sites. DNA fragments provided in pUC57-Amp were digested by EcoRI and SalI and were then cloned into EcoRI/SalI-digested pBAD30 (39).

Chimeric genes *CmE* and *EmC* were constructed using the gene splicing by overlap extension technique (40). *CmE* fragments were PCR amplified with primer pair 94vcrx145.f/eexC182.r and primer pair eexC280.f/pAsa4cTrrA.r using pVCR94^{5p} and pAsa4c as the templates, respectively. *EmC* fragments were PCR amplified with primer pair ptrxA2.f/eexC279.r and primer pair eexC183.f/94vcrx145.rev using pAsa4c and pVCR94^{5p} as the templates, respectively. After extension and PCR amplification, the DNA fragments containing the chimeric genes were digested with EcoRI and cloned into EcoRI-digested pBAD30.

All constructions were verified by sequencing performed by the Plateforme de Séquençage et de Génotypage du Centre de Recherche du CHUL (Québec, QC, Canada).

Phylogenetic analyses. Primary sequences of Eex and TraG homologs were obtained by using the NCBI's tblastn algorithm against the nr/nt database restricted to *Gammaproteobacteria* (taxid: 1236). Only primary sequences sharing more than 40% identity with EexC and 50% identity with TraG_C over 80% minimum coverage were included in subsequent analyses. Since the SGI1 TraG_S and SXT TraG proteins do not fit this criterion (37% identity with TraG_C over 99% coverage [27, 41]), both were included manually in the TraG data set as an outgroup. Eex and TraG primary sequences were first clustered with CD-HIT (42) (sequence identity cutoff of 1) prior to alignment. Exhaustive data on the composition of these clusters are available in Data Set S1 in the supplemental material.

Evolutionary analyses of Eex or TraG proteins were performed within MEGA7 (43), and data were inferred by using the maximum likelihood (PhyML) (44) method based on the JTT matrix-based (Eex) and Whelan and Goldman (WAG) + Freq. (TraG) models (45, 46). Phylogenetic analyses were computed using amino acid alignments generated by MUSCLE (47). Poorly aligned regions were removed using trimAl v1.3 software and the automated heuristic approach (48) prior to phylogenetic analyses. The initial tree(s) for the heuristic search was obtained automatically by applying neighbor-joining and BioNJ algorithms to a matrix of pairwise distances estimated using a JTT model and then selecting the topology with the superior log likelihood value. A discrete gamma distribution was used to model evolutionary rate differences among sites (5 categories [+G, parameter = 3.0102]) for the TraG phylogenetic tree.

In silico typing of plasmids. Replicon types of plasmids listed in Data Sets S1 and S2 were determined using PlasmidFinder 2.0 (49) and the *Enterobacteriaceae* database (95% identity, 60% coverage).

Protein secondary structure prediction. The secondary structures of EexE and EexC were predicted using PSIPRED protein sequence analysis workbench v3.3 (50). Subcellular localization was predicted using PSORT-B v3.0 (51). Transmembrane protein topology was predicted using MEMSAT3 (52). Comparisons of the aligned Kyte-Doolittle hydropathy plots of EexC and EexE were generated using AlignMe with the Fast method (53).

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at <https://doi.org/10.1128/JB.00731-18>.

SUPPLEMENTAL FILE 1, XLSX file, 0.1 MB.

ACKNOWLEDGMENTS

We thank Steve Charette from Université Laval for the kind gift of pAsa4c. We are grateful to Nicolas Carraro and Kevin Neil for technical assistance and Alain Lavigne for critical reading of the manuscript.

This work was supported by a Discovery Grant (2016-04365) from the Natural Sciences and Engineering Council of Canada (NSERC) and a Project Grant (PJT-153071) from the Canadian Institutes of Health Research (CIHR) to V.B. K.T.H. was supported by a postdoctoral fellowship (SPE20170336797) from the Fondation pour la Recherche Médicale (FRM, France).

The funding agencies had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

REFERENCES

- Harmer CJ, Hall RM. 2015. The A to Z of A/C plasmids. *Plasmid* 80:63–82. <https://doi.org/10.1016/j.plasmid.2015.04.003>.
- Weill F-X, Dommann D, Njamkepo E, Tarr C, Rauzier J, Fawal N, Keddy KH, Salje H, Moore S, Mukhopadhyay AK, Bercion R, Luquero FJ, Ngandjio A, Dosso M, Monakhova E, Garin B, Bouchier C, Pazzani C, Mutreja A, Grunow R, Sidikou F, Bonte L, Breurec S, Damian M, Njanpop-Lafourcade B-M, Sapriel G, Page A-L, Hamze M, Henkens M, Chowdhury G, Mengel M, Koeck J-L, Fournier J-M, Dougan G, Grimont PAD, Parkhill J, Holt KE, Piarroux R, Ramamurthy T, Quilici M-L, Thomson NR. 2017. Genomic history of the seventh pandemic of cholera in Africa. *Science* 358: 785–789. <https://doi.org/10.1126/science.aad5901>.
- Miller RA, Harbottle H. 2018. Antimicrobial drug resistance in fish pathogens. *Microbiol Spectr* 6. <https://doi.org/10.1128/microbiolspec>.
- Rozwandowicz M, Brouwer MSM, Fischer J, Wagenaar JA, Gonzalez-Zorn

- B, Guerra B, Mevius DJ, Hordijk J. 2018. Plasmids carrying antimicrobial resistance genes in Enterobacteriaceae. *J Antimicrob Chemother* 73: 1121–1137. <https://doi.org/10.1093/jac/dkx488>.
5. Doublet B, Boyd D, Mulvey MR, Cloeckaert A. 2005. The *Salmonella* genomic island 1 is an integrative mobilizable element. *Mol Microbiol* 55:1911–1924. <https://doi.org/10.1111/j.1365-2958.2005.04520.x>.
6. Carraro N, Rivard N, Ceccarelli D, Colwell RR, Burrus V. 2016. IncA/C conjugative plasmids mobilize a new family of multidrug resistance islands in clinical *Vibrio cholerae* non-O1/non-O139 isolates from Haiti. *mBio* 7:e00509-16. <https://doi.org/10.1128/mBio.00509-16>.
7. Novick RP. 1969. Extrachromosomal inheritance in bacteria. *Bacteriol Rev* 33:210–263.
8. Chandran Darbari V, Waksman G. 2015. Structural biology of bacterial type IV secretion systems. *Annu Rev Biochem* 84:603–629. <https://doi.org/10.1146/annurev-biochem-062911-102821>.
9. Arutyunov D, Frost LS. 2013. F conjugation: back to the beginning. *Plasmid* 70:18–32. <https://doi.org/10.1016/j.plasmid.2013.03.010>.
10. Audette GF, Manchak J, Beatty P, Klimke WA, Frost LS. 2007. Entry exclusion in F-like plasmids requires intact TraG in the donor that recognizes its cognate TraS in the recipient. *Microbiology* 153:442–451. <https://doi.org/10.1099/mic.0.2006/001917-0>.
11. Gunton JE, Ussher JER, Rooker MM, Wetsch NM, Alonso G, Taylor DE. 2008. Entry exclusion in the IncHI1 plasmid R27 is mediated by EexA and EexB. *Plasmid* 59:86–101. <https://doi.org/10.1016/j.plasmid.2007.11.004>.
12. Sakuma T, Tazumi S, Furuya N, Komano T. 2013. ExcA proteins of IncI1 plasmid R64 and IncIy plasmid R621a recognize different segments of their cognate TraY proteins in entry exclusion. *Plasmid* 69:138–145. <https://doi.org/10.1016/j.plasmid.2012.11.004>.
13. Pohlman RF, Genetti HD, Winans SC. 1994. Entry exclusion of the IncN plasmid pKM101 is mediated by a single hydrophilic protein containing a lipid attachment motif. *Plasmid* 31:158–165. <https://doi.org/10.1006/plas.1994.1017>.
14. Bolland S, Lloa M, Avila P, de la Cruz F. 1990. General organization of the conjugal transfer genes of the IncW plasmid R388 and interactions between R388 and IncN and IncP plasmids. *J Bacteriol* 172:5795–5802. <https://doi.org/10.1128/jb.172.10.5795-5802.1990>.
15. Haase J, Kalkum M, Lanka E. 1996. TrbK, a small cytoplasmic membrane lipoprotein, functions in entry exclusion of the IncP alpha plasmid RP4. *J Bacteriol* 178:6720–6729. <https://doi.org/10.1128/jb.178.23.6720-6729.1996>.
16. Marrero J, Waldor MK. 2005. Interactions between inner membrane proteins in donor and recipient cells limit conjugal DNA transfer. *Dev Cell* 8:963–970. <https://doi.org/10.1016/j.devcel.2005.05.004>.
17. Christie PJ. 4 August 2016, posting date. The mosaic type IV secretion systems. *EcoSal Plus* 2016. <https://doi.org/10.1128/ecosalplus.ESP-0020-2015>.
18. Anthony KG, Klimke WA, Manchak J, Frost LS. 1999. Comparison of proteins involved in pilus synthesis and mating pair stabilization from the related plasmids F and R100-1: insights into the mechanism of conjugation. *J Bacteriol* 181:5149–5159.
19. Marrero J, Waldor MK. 2007. The SXT/R391 family of integrative conjugative elements is composed of two exclusion groups. *J Bacteriol* 189: 3302–3305. <https://doi.org/10.1128/JB.01902-06>.
20. Garcillán-Barcia MP, de la Cruz F. 2008. Why is entry exclusion an essential feature of conjugative plasmids? *Plasmid* 60:1–18. <https://doi.org/10.1016/j.plasmid.2008.03.002>.
21. Fricke WF, Welch TJ, McDermott PF, Mammel MK, LeClerc JE, White DG, Cebula TA, Ravel J. 2009. Comparative genomics of the IncA/C multidrug resistance plasmid family. *J Bacteriol* 191:4750–4757. <https://doi.org/10.1128/JB.00189-09>.
22. Datta N, Hedges RW. 1973. R factors of compatibility group A. *J Gen Microbiol* 74:335–337. <https://doi.org/10.1099/00221287-74-2-335>.
23. Hedges RW, Datta N. 1971. R factors giving chloramphenicol resistance. *Nature* 234:220. <https://doi.org/10.1038/234220b0>.
24. Datta N, Hedges RW. 1972. R factors identified in Paris, some conferring gentamicin resistance, constitute a new compatibility group. *Ann Inst Pasteur (Paris)* 123:849–852.
25. Hedges RW. 1974. R factors from Providence. *J Gen Microbiol* 81:171–181. <https://doi.org/10.1099/00221287-81-1-171>.
26. Ambrose SJ, Harmer CJ, Hall RM. 2018. Compatibility and entry exclusion of IncA and IncC plasmids revisited: IncA and IncC plasmids are compatible. *Plasmid* 96–97:7–12. <https://doi.org/10.1016/j.plasmid.2018.02.002>.
27. Carraro N, Durand R, Rivard N, Anquetil C, Barrette C, Humbert M, Burrus V. 2017. *Salmonella* genomic island 1 (SGI1) reshapes the mating apparatus of IncC conjugative plasmids to promote self-propagation. *PLoS Genet* 13:e1006705. <https://doi.org/10.1371/journal.pgen.1006705>.
28. Douard G, Praud K, Cloeckaert A, Doublet B. 2010. The *Salmonella* genomic island 1 is specifically mobilized in trans by the IncA/C multidrug resistance plasmid family. *PLoS One* 5:e15302. <https://doi.org/10.1371/journal.pone.0015302>.
29. Carraro N, Matteau D, Luo P, Rodrigue S, Burrus V. 2014. The master activator of IncA/C conjugative plasmids stimulates genomic islands and multidrug resistance dissemination. *PLoS Genet* 10:e1004714. <https://doi.org/10.1371/journal.pgen.1004714>.
30. Lawley TD, Klimke WA, Gubbins MJ, Frost LS. 2003. F factor conjugation is a true type IV secretion system. *FEMS Microbiol Lett* 224:1–15. [https://doi.org/10.1016/S0378-1097\(03\)00430-0](https://doi.org/10.1016/S0378-1097(03)00430-0).
31. Ambrose SJ, Harmer CJ, Hall RM. 2018. Evolution and typing of IncC plasmids contributing to antibiotic resistance in Gram-negative bacteria. *Plasmid* 99:40–55. <https://doi.org/10.1016/j.plasmid.2018.08.001>.
32. Marrero J, Waldor MK. 2007. Determinants of entry exclusion within Eex and TraG are cytoplasmic. *J Bacteriol* 189:6469–6473. <https://doi.org/10.1128/JB.00522-07>.
33. Bioteau A, Durand R, Burrus V. 2018. Redefinition and unification of the SXT/R391 family of integrative and conjugative elements. *Appl Environ Microbiol* 84:e00485-18. <https://doi.org/10.1128/AEM.00485-18>.
34. Spagnoletti M, Ceccarelli D, Rieux A, Fondi M, Taviani E, Fani R, Colombo MM, Colwell RR, Balloux F. 2014. Acquisition and evolution of SXT-R391 integrative conjugative elements in the seventh-pandemic *Vibrio cholerae* lineage. *mBio* 5:e01356-14. <https://doi.org/10.1128/mBio.01356-14>.
35. Achtman M, Kennedy N, Skurray R. 1977. Cell-cell interactions in conjugating *Escherichia coli*: role of traT protein in surface exclusion. *Proc Natl Acad Sci U S A* 74:5104–5108. <https://doi.org/10.1073/pnas.74.11.5104>.
36. Dower WJ, Miller JF, Ragsdale CW. 1988. High efficiency transformation of *E. coli* by high voltage electroporation. *Nucleic Acids Res* 16: 6127–6145. <https://doi.org/10.1093/nar/16.13.6127>.
37. Datsenko KA, Wanner BL. 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A* 97:6640–6645. <https://doi.org/10.1073/pnas.120163297>.
38. St-Pierre F, Cui L, Priest DG, Endy D, Dodd IB, Shearwin KE. 2013. One-step cloning and chromosomal integration of DNA. *ACS Synth Biol* 2:537–541. <https://doi.org/10.1021/sb400021j>.
39. Guzman LM, Belin D, Carson MJ, Beckwith J. 1995. Tight regulation, modulation, and high-level expression by vectors containing the arabinose *PBAD* promoter. *J Bacteriol* 177:4121–4130. <https://doi.org/10.1128/jb.177.14.4121-4130.1995>.
40. Horton RM, Hunt HD, Ho SN, Pullen JK, Pease LR. 1989. Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension. *Gene* 77:61–68. [https://doi.org/10.1016/0378-1119\(89\)90359-4](https://doi.org/10.1016/0378-1119(89)90359-4).
41. Wozniak RAF, Fouts DE, Spagnoletti M, Colombo MM, Ceccarelli D, Garriss G, Dery C, Burrus V, Waldor MK. 2009. Comparative ICE genomics: insights into the evolution of the SXT/R391 family of ICEs. *PLoS Genet* 5:e1000786. <https://doi.org/10.1371/journal.pgen.1000786>.
42. Li W, Godzik A. 2006. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* 22: 1658–1659. <https://doi.org/10.1093/bioinformatics/btl158>.
43. Kumar S, Stecher G, Tamura K. 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol* 33: 1870–1874. <https://doi.org/10.1093/molbev/msw054>.
44. Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol* 59: 307–321. <https://doi.org/10.1093/sysbio/syq010>.
45. Jones DT, Taylor WR, Thornton JM. 1992. The rapid generation of mutation data matrices from protein sequences. *Comput Appl Biosci* 8:275–282.
46. Whelan S, Goldman N. 2001. A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. *Mol Biol Evol* 18:691–699. <https://doi.org/10.1093/oxfordjournals.molbev.a003851>.
47. Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32:1792–1797. <https://doi.org/10.1093/nar/gkh340>.
48. Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinforma Oxf Engl* 25:1972–1973. <https://doi.org/10.1093/bioinformatics/btp348>.
49. Carattoli A, Zankari E, García-Fernández A, Voldby Larsen M, Lund O, Villa L, Møller Aarestrup F, Hasman H. 2014. In silico detection and typing of

- plasmids using PlasmidFinder and plasmid multilocus sequence typing. *Antimicrob Agents Chemother* 58:3895–3903. <https://doi.org/10.1128/AAC.02412-14>.
50. Jones DT. 1999. Protein secondary structure prediction based on position-specific scoring matrices. *J Mol Biol* 292:195–202. <https://doi.org/10.1006/jmbi.1999.3091>.
 51. Gardy JL, Spencer C, Wang K, Ester M, Tusn  dy GE, Simon I, Hua S, deFays K, Lambert C, Nakai K, Brinkman F. 2003. PSORT-B: improving protein subcellular localization prediction for Gram-negative bacteria. *Nucleic Acids Res* 31:3613–3617. <https://doi.org/10.1093/nar/gkg602>.
 52. Jones DT. 2007. Improving the accuracy of transmembrane protein topology prediction using evolutionary information. *Bioinforma Oxf Engl* 23:538–544. <https://doi.org/10.1093/bioinformatics/btl677>.
 53. Stamm M, Staritzbichler R, Khafizov K, Forrest LR. 2014. AlignMe—a membrane protein sequence alignment web server. *Nucleic Acids Res* 42:W246–W251. <https://doi.org/10.1093/nar/gku291>.
 54. Singer M, Baker TA, Schnitzler G, Deischel SM, Goel M, Dove W, Jaacks KJ, Grossman AD, Erickson JW, Gross CA. 1989. A collection of strains containing genetically linked alternating antibiotic resistance elements for genetic mapping of *Escherichia coli*. *Microbiol Rev* 53:1–24.
 55. Grenier F, Matteau D, Baby V, Rodrigue S. 2014. Complete genome sequence of *Escherichia coli* BW25113. *Genome Announc* 2:e01038-14. <https://doi.org/10.1128/genomeA.01038-14>.
 56. Carraro N, Sauv   M, Matteau D, Lauzon G, Rodrigue S, Burrus V. 2014. Development of pVCR94  X from *Vibrio cholerae*, a prototype for studying multidrug resistant IncA/C conjugative plasmids. *Front Microbiol* 5:44. <https://doi.org/10.3389/fmicb.2014.00044>.
 57. Ceccarelli D, Daccord A, Ren   M, Burrus V. 2008. Identification of the origin of transfer (*oriT*) and a new gene required for mobilization of the SXT/R391 family of integrating conjugative elements. *J Bacteriol* 190:5328–5338. <https://doi.org/10.1128/JB.00150-08>.
 58. Tanaka KH, Vincent AT, Trudel MV, Paquet VE, Frenette M, Charette SJ. 2016. The mosaic architecture of *Aeromonas salmonicida* subsp. *salmonicida* pAsa4 plasmid and its consequences on antibiotic resistance. *PeerJ* 4:e2595. <https://doi.org/10.7717/peerj.2595>.
 59. Carraro N, Sauv   M, Matteau D, Lauzon G, Rodrigue S, Burrus V. 6 February 2014. Development of pVCR94  X from *Vibrio cholerae*, a prototype for studying multidrug resistant IncA/C conjugative plasmids. *Front Microbiol*. <https://doi.org/10.3389/fmicb.2014.00044>.
 60. Cabez  n E, Sastre JI, de la Cruz F. 1997. Genetic evidence of a coupling role for the TraG protein family in bacterial conjugation. *Mol Gen Genet* 254:400–406.
 61. Datta S, Costantino N, Court DL. 2006. A set of recombinering plasmids for gram-negative bacteria. *Gene* 379:109–115. <https://doi.org/10.1016/j.gene.2006.04.018>.
 62. Matteau D, Pepin M-E, Baby V, Gauthier S, Arango Giraldo M, Knight TF, Rodrigue S. 17 March 2017. Development of *oriC*-based plasmids for *Mesoplasma florum*. *Appl Environ Microbiol*. <https://doi.org/10.1128/AEM.03374-16>.

CHAPITRE 3

RÉSULTATS COMPLÉMENTAIRES

3.1 Matériel et méthodes

3.1.1. Culture bactérienne

Les souches bactériennes, les plasmides et les îlots génomiques utilisés dans ce chapitre sont décrits dans le **Tableau 1**.

Les conditions de culture et de conservation des souches, ainsi que les concentrations d'antibiotiques utilisées sont identiques à celles listées dans l'article (2.2 Article : Entry Exclusion of Conjugative Plasmids of the IncA, IncC, and Related Untyped Incompatibility Groups, page 36).

Souche, plasmide ou îlot génomique	Génotype ou phénotype	Source ou référence
<i>Souches</i>		
CAG18439	MG1655 <i>lacZU118 lacI42::Tn10</i> (Tc)	(Singer <i>et al.</i> , 1989)
BW25113	F ⁻ Δ (<i>araD-araB</i>)567, Δ <i>lacZ4787</i> (:: <i>rrnB-3</i>), λ , <i>rph-1</i> , Δ (<i>rhaD-rhaB</i>)568, <i>hsdR514</i>	(Datsenko et Wanner, 2000)
BW25113Nx	Dérivé Nx ^R de BW25113 (GG56) (Nx)	(Grenier <i>et al.</i> , 2014)
MG1655Rf	Dérivé Rf ^R de MG1655 (VB112) (Rf)	(Ceccarelli <i>et al.</i> , 2008)
<i>Plasmides</i>		
pVCR94	Plasmide conjugatif IncC de <i>V. cholerae</i> O1 El Tor (Su Tm Cm Ap Tc Sm)	(Carraro <i>et al.</i> , 2014b)
pVCR94 Δ X	Dérivé Su ^R de pVCR94, sans le grand fragment contenant le cluster de multirésistance aux antibiotique (Su)	(Carraro <i>et al.</i> , 2014b)
pVCR94 ^{Sp}	Dérivé Sp ^R de pVCR94 Δ X (pVCR94 Δ X2) (Sp Su)	(Carraro <i>et al.</i> , 2014a)
pVCR94 ^{Kn}	Dérivé Kn ^R de pVCR94 Δ X (pVCR94 Δ X3) (Kn Su)	(Carraro <i>et al.</i> , 2014a)
pVCR94 ^{Sp} Δ eexC	pVCR94 ^{Sp} Δ eexC (Sp Su)	(Humbert <i>et al.</i> , 2019)
pVCR94 ^{Sp} Δ eex:: <i>cat</i>	pVCR94 ^{Sp} Δ eex:: <i>cat</i> (Sp Su Cm)	Cette étude
pVCR94 ^{Sp} Δ vcrx085	pVCR94 ^{Sp} Δ vcrx085 (Sp Su)	Cette étude
pVCR94 ^{Sp} Δ eexC Δ vcrx085	pVCR94 ^{Sp} Δ eexC Δ vcrx085 (Sp Su)	Cette étude
pVCR94 ^{Sp} Δ traG	pVCR94 ^{Sp} Δ traG (Sp Su)	(Carraro <i>et al.</i> , 2017a)
pBAD30	<i>ori</i> _{p15A} <i>araC</i> P _{BAD} (Ap)	(Guzman <i>et al.</i> , 1995)
peexC	pBAD30:: <i>eexC</i> (<i>vcrx145</i> of pVCR94 Δ X) (Ap)	(Humbert <i>et al.</i> , 2019)

<i>peexE</i>	pBAD30:: <i>eexE</i> (sequence 156,821 to 157,174 of pAsa4c) (Ap)	(Humbert <i>et al.</i> , 2019)
<i>ptragE</i>	pBAD30:: <i>traGE</i> (sequence 153,216 to 156,770 of pAsa4c) (Ap)	(Humbert <i>et al.</i> , 2019)
<i>pCmE</i>	pBAD30:: <i>eexCmE</i> (Chimère 1) (Ap)	(Humbert <i>et al.</i> , 2019)
<i>pEmC</i>	pBAD30:: <i>eexEmC</i> (Chimère 2) (Ap)	(Humbert <i>et al.</i> , 2019)
<i>peexCEch3</i>	pBAD30 contenant la chimère <i>eexC-eexE</i> n°3 (Ap)	Cette étude
<i>peexCEch4</i>	pBAD30 contenant la chimère <i>eexC-eexE</i> n°4 (Ap)	Cette étude
<i>peexCEch5</i>	pBAD30 contenant la chimère <i>eexC-eexE</i> n°5 (Ap)	Cette étude
<i>peexCEch6</i>	pBAD30 contenant la chimère <i>eexC-eexE</i> n°6 (Ap)	Cette étude
<i>peexCEch7</i>	pBAD30 contenant la chimère <i>eexC-eexE</i> n°7 (Ap)	Cette étude
<i>peexCEch8</i>	pBAD30 contenant la chimère <i>eexC-eexE</i> n°8 (Ap)	Cette étude
<i>peexCEch9</i>	pBAD30 contenant la chimère <i>eexC-eexE</i> n°9 (Ap)	Cette étude
<i>peexCEch10</i>	pBAD30 contenant la chimère <i>eexC-eexE</i> n°10 (Ap)	Cette étude
<i>ps009</i>	pBAD30:: <i>s009</i> (Ap)	Cette étude
<i>pvcrx085</i>	pBAD30:: <i>vcrx085</i> (Ap)	Cette étude
<i>pClo</i>	CloDF13::TnAΔEcoRV (pSU4628) (Ap)	(Cabezón <i>et al.</i> , 1997)
<i>pSIM6</i>	Expression thermo-inductible de la recombinaison λRed (<i>ori^{ts}</i> Ap)	(Datta <i>et al.</i> , 2006)
<i>pKD3</i>	Matrice Cm ^R pour les délétions	(Datsenko et Wanner, 2000)
<i>pKD4</i>	Matrice Kn ^R pour les délétions	(Datsenko et Wanner, 2000)
<i>pCP20</i>	Expression de la recombinase FLP (<i>ori^{ts}</i> Ap)	(Cherepanov et Wackernagel, 1995)
<i>Îlots génomiques</i>		
SGI1	SGI1 inséré en 3' de <i>trmE</i> (Ap Cm Sp Sm Su Tc)	(Carraro <i>et al.</i> , 2014a)
SGI1 ^{Kn}	Mutant ΔIn104:: <i>aph</i> de SGI1 sans l'intégron In104 (Kn)	(Carraro <i>et al.</i> , 2017a)
SGI1 ^{Cm}	Mutant ΔIn104:: <i>cat</i> de SGI1 sans l'intégron In104 (Cm)	Cette étude
SGI1 ^{Kn} Δs009	SGI1 ^{Kn} Δs009 (Kn)	Cette étude

Tableau 1. Souches d'*E. coli* dérivées de K-12, plasmides et îlots génomiques utilisés

Ap, ampicilline ; Cm, chloramphénicol ; Kn, kanamycine ; Nx, acide nalidixique ; Rf, rifampicine ; Sp, spectinomycine ; Sm, streptomycine ; Su, sulfamethoxazole ; Tc, tétracycline ; Tm, triméthoprime ; ts, thermosensible

3.1.2. Essais de conjugaison

Les essais de conjugaison ont été effectués en suivant le protocole présenté dans l'article (2.2 Article : Entry Exclusion of Conjugative Plasmids of the IncA, IncC, and Related Untyped Incompatibility Groups, page 36). Les transferts conjugatifs dont les résultats sont présentés dans la figure 10A ont été effectués en utilisant MG1655Rf comme souche donneuse et

BW25113Nx comme souche réceptrice. Tous les autres transferts présentés dans ce chapitre ont été effectués en utilisant BW25113Nx comme donneuse et CAG18439 comme réceptrice.

3.1.3. Méthodes de biologie moléculaire

Les méthodes de biologie moléculaire, telles que l'extraction et la purification d'ADN, les PCR et l'électroporation, sont identiques à celles présentées dans l'article (2.2 Article : Entry Exclusion of Conjugative Plasmids of the IncA, IncC, and Related Untyped Incompatibility Groups, page 36).

3.1.4. Construction des plasmides et des souches

Les souches bactériennes, les plasmides et les îlots génomiques utilisés dans ce chapitre sont décrits dans le **Tableau 1**. Les amorces utilisées pour leur construction sont listées dans le **Tableau 2**.

Les plasmides encodant les chimères de EexC et EexE 3 à 10 ont été construits en utilisant le kit « Q5 site-directed mutagenesis » de NEB. Conformément aux instructions de ce kit, le design des amorces a été effectué avec NEBaseChanger. Les chimères 3, 5, 7 et 9 ont été construites en utilisant *peexC* comme matrice et les couples d'amorces *eexCmutv4.f/eexCmutv4.r*, *eexCmutv5.f/eexCmutv5.r*, *eexCmLA.f/eexCmLA.r* et *eexCmSA.f/eexCmSA.r*, respectivement. Les chimères 4, 6, 8 et 10 ont été construites en utilisant *peexA* comme matrice et les couples d'amorces *eexAmutv4.f/eexAmutv4.r*, *eexAmutv5.f/eexAmutv5.r*, *eexAmLC.f/eexAmLC.r* et *eexAmSC.f/eexAmSC.r*, respectivement.

Le plasmide *ps009* a été obtenue en amplifiant le gène *s009* par PCR en utilisant les amorces SGI1S009EcoRI.for/SGI1S009EcoRI.rev, contenant des sites EcoRI, et une préparation d'ADN génomique contenant SGI1^{Kn} comme matrice. Le fragment amplifié a ensuite été digéré par EcoRI puis cloné dans un plasmide pBAD30 lui aussi digéré par EcoRI.

SGI1^{Cm} a été construit en utilisant la technique mise au point par Datsenko et Wanner (2000) (one-step chromosomal gene inactivation) avec le plasmide pSIM6. La cassette de résistance au chloramphénicol a été amplifiée avec les amorces SGI1In104cm2.f/SGI1In104cm2.r à partir du plasmide pKD3. La délétion de *s009* dans SGI1^{Kn} et de *vcx085* dans pVCR94^{Sp} et pVCR94^{Sp}Δ*eexC* ont été effectuées avec la même méthode. Pour SGI1^{Kn}Δ*s009*, la cassette de résistance a été amplifiée par les amorces SGI1del009.for/SGI1del009.rev à partir de pKD4. Pour pVCR94^{Sp}Δ*vcx085* et pVCR94^{Sp}Δ*eexC*Δ*vcx085*, les amorces Fw-Dltvcx085/Rv-Dltvcx085 et le plasmide pKD4 ont été utilisés. Les cassettes de résistances ont ensuite été enlevées en introduisant le plasmide pCP20 par électroporation dans les souches obtenues, conformément à la technique mise au point par Datsenko et Wanner (2000). Ces délétions ont été vérifiées par PCR et par l'analyse des résistances aux antibiotiques.

Lorsque nécessaire, ces constructions ont été vérifiées par séquençage effectué par la Plateforme de Séquençage et de Génotypage du Centre de Recherche du CHUL (Québec, QC, Canada).

Nom de l'amorce	Séquence nucléotidique (de 5' à 3') ^a
eexCmutv4.f	ggctccaagtcgaaaACTGGTGGGTTTCAAGATTC
eexCmutv4.r	acgaagatccacgccCTCTTTATATGCTGACCG
eexCmutv5.f	agctcgttgGGAGACGAATAAGAATTTCG
eexCmutv5.r	ttttcaactagGTCAAAATCATAGGAATCTTG
eexCmLA.f	tcgttggggtgaaTAAGAATTCGAGCTCGGTAC
eexCmLA.r	gctttttcaactagATCATAGGAATCTTGAAACCC
eexCmSA.f	gaaaaaTTCCATATTGGAGCTGGAG
eexCmSA.r	aactagATCATAGGAATCTTGAAACCC
eexAmutv4.f	cgcagaaaaaaacaaACGATTGATGGCGTCAATG
eexAmutv4.r	ccgggtaggtggtgcCTCTCTTCTCGCCTTTATG
eexAmutv5.f	tattggagctGGTGAATGAGAATTCGAG
eexAmutv5.r	tggaagcgctgGTCAATAGAAAGAGGACC
eexAmLC.f	attggagctggagacgaaTGAGAATTCGAGCTCGGT
eexAmLC.r	atggaagcgctggtcaaaGTCAATAGAAAGAGGACCATTG
eexAmSC.f	cagcgcGCTCGTTGGGGTGAATGA
eexAmSC.r	gtcaaaGTCAATAGAAAGAGGACCATTG
SGI1S009EcoRI.for	NNNNNNGAATTCAAGGAGGAATAATAAATGAAAAAGGCGCTATTCTCTTTGT
SGI1S009EcoRI.rev	NNNNNNGAATTCTCAGGAGCCTTTGTTTTTATTCAAT
SGI1In104cm2.f	TATTTTAACATGTATATAGATTAAATTCCAATCACTTGTCGGAATAGGAACCTTCATTTA

SGI1In104cm2.r	CAAACCACCATTTTCTTAACAGTTCATAACTTAGCAAGTTGCCTACCTGTGACGGAAGAT
SGI1del009.for	GATAATCTATAATCGGTTACATATTAATCAGGAGGCGTAAgtgtaggctggagctgcttc
SGI1del009.rev	TTAACGTTAAGTCTATAATGAATAAAAAACAAAGGCTCTCATTACATATGAATATCCTCCTTA
Fw-Dltvcx085	AAAAAGGGGCCTTTTCGGCCCCTTTCTCATGTAGATTGAGTGTAGGCTGGAGCTGCTTCG
Rv-Dltvcrx085	GTAAATCAGATTCAACTCAGGTAAGAACGGAGAACAACCATATGAATATCCTCCTTA

Tableau 2. Séquence nucléotidique des amorces utilisées

^a Les sites de restriction EcoRI sont soulignés

3.1.5. Mutagenèse par insertion de transposons

Le protocole de mutagenèse par insertion de transposons est identique à celui présenté par Roy *et al.* (2020). Il a été effectué sur la souche BW25113Nx contenant pVCR94^{Kn} Δ eex::cat. Le transfert du plasmide contenant le transposon a été effectué avec un ratio donneuses/réceptrices de 5:1 et deux heures d'incubation à 37°C. Les réceptrices dans lesquelles Tn5 est intégré ont été sélectionnées avec de l'acide nalidixique, de la kanamycine, du chloramphénicol et de la spectinomycine. Le transfert de pClo vers ces réceptrices, a été effectué avec 25 ml de cellules donneuses (MG1655Rf pVCR94^{Sp}) et 25 ml de cellule réceptrice et une incubation de deux heures à 37°C. Les transconjugants ayant reçu pClo ont été sélectionné avec de l'ampicilline, de l'acide nalidixique, de la kanamycine, du chloramphénicol et de la spectinomycine. La préparation des banques, le séquençage et l'analyse des données de HDTM ont été effectué d'après le protocole de Roy *et al.* (2020). Les données de l'HDTM ont été alignées sur la séquence de pVCR94^{Kn} Δ eexC Δ acr2.

3.1.6. Analyse phylogénétique

Les séquences nucléotidiques correspondant à des homologues de Vcrx085 ont été obtenues en utilisant l'algorithme tBLASTn de NCBI (Altschul *et al.*, 1990) avec le groupe des Gammaprotéobactéries (taxid : 1236) comme cible. Seules les séquences partageant plus de 30% d'identité sur une couverture de 90% minimum ont été incluses dans les analyses suivantes. Ces séquences ont été traduites en séquences protéiques par Transeq (Madeira *et al.*, 2019)

(codon : bactérien) puis classées en clusters de séquences identiques par CD-HIT (Li et Godzik, 2006). La composition précise de ces clusters est disponible dans le **Tableau S1** (page 76). Un représentant de chaque cluster a été conservé pour la suite. Un alignement des séquences a été obtenu avec MUSCLE (Edgar, 2004), et les régions faiblement alignées ont été enlevées par le logiciel trimAl v1.3 (Capella-Gutiérrez *et al.*, 2009) en utilisant l'approche heuristique automatique (appelée « automated 1 »). L'arbre phylogénétique a été obtenu avec MEGA7 (Kumar *et al.*, 2016) en utilisant la méthode PhyML (Guindon *et al.*, 2010) basée sur le modèle LG + G + I.

3.1.7. Prédiction de structure secondaire de protéines

La localisation de Vcrx085 a été prédite par PSORT-b (Gardy *et al.*, 2003). La présence d'un peptide signal a été analysée par SignalP-5.0 (Almagro Armenteros *et al.*, 2019) et la structure secondaire a été prédite par PSI-PRED (Jones, 1999).

3.2 Résultats

3.2.1. Exclusion d'entrée

3.2.1.1 Chimères EexC-EexE

Dans l'article présenté dans le chapitre 2, nous avons pu identifier la partie de Eex permettant la spécificité des groupes d'exclusion : il s'agit des 40 acides aminés à l'extrémité C-terminale de la protéine (**figure 5B**, chimères 1 et 2). J'ai tenté par la suite d'identifier plus précisément les acides aminés qui permettaient la reconnaissance entre Eex et TraG.

Pour cela, j'ai entrepris d'échanger des portions plus petites de EexC et EexE, en choisissant des régions contenant des différences importantes d'acides aminés. J'ai effectué un alignement des séquences protéiques avec MUSCLE (Edgar, 2004) qui indique à la fois les acides aminés

identiques entre les deux protéines mais aussi ceux qui ont des fonctions similaires (**figure 5A**). En me basant sur cet alignement, j'ai choisi les régions à échanger et construit les gènes chimériques par mutagenèse dirigée des plasmides *peexC* et *peexE*. Les plasmides contenant les gènes chimériques ont ensuite été introduits dans des bactéries vers lesquelles j'ai testé le transfert de $pVCR94^{Sp}$ et de $pVCR94^{Sp}\Delta traG$ complémenté par *ptragE*.

La **figure 5B** présente ces différentes chimères et leur faculté à exclure $pVCR94$ ou $pAsa4c$ (essais réalisés avec $pVCR94^{Sp}\Delta traG$ complémenté par *ptragE*). Les deux premières chimères correspondent à celles qui ont été employées précédemment dans l'article (2.2 Article : Entry Exclusion of Conjugative Plasmids of the IncA, IncC, and Related Untyped Incompatibility Groups, page 36). L'échange des 10 acides aminés à la position 90 (**figure 5B**, chimères 3 et 4) n'a eu aucun impact sur la spécificité de l'exclusion exercée par *EexC* ou *EexE*. En revanche, de nombreux essais d'échange d'acides aminés situés entre les positions 110 et 120 ont mené à une perte totale de l'exclusion des chimères (**figure 5B**, chimères 5 à 10). On peut donc supposer que la région contenant les 10 à 20 derniers acides aminés de *EexC* et *EexE* présente une structure trop différente pour permettre la construction de chimères fonctionnelles. Cette région est essentielle pour permettre à *Eex* d'exercer une exclusion d'entrée, et c'est probablement aussi dans cette région que se situent les acides aminés permettant la spécificité entre *Eex* et *TraG*.

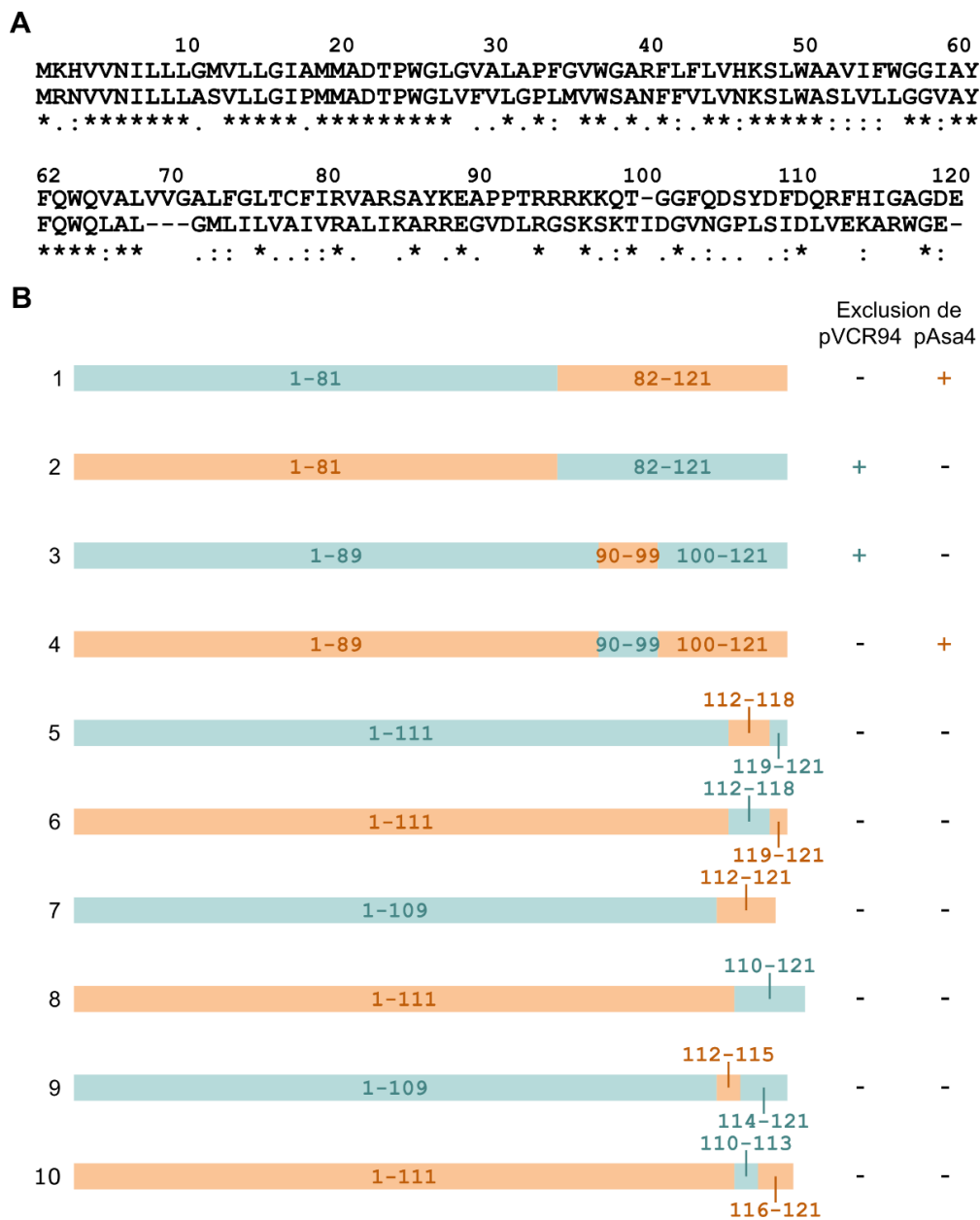


Figure 5. Effet des chimères EexC-EexE sur l'exclusion

A : alignement des séquences protéiques de EexC et EexE obtenu par MUSCLE, acides aminés identiques (*), fortement similaires (:), faiblement similaires (.).

B : représentation des différentes chimères construites (portions provenant de EexC en bleu et portions provenant de EexE en orange) et leur capacité à exclure pVCR94 ou pAsa4c (+ : exclusion, - : pas d'exclusion)

3.2.1.2 Exclusion d'entrée de SGI1

Pour savoir si SGI1 possède ou non un système d'exclusion, j'ai effectué le transfert de SGI1^{Cm}, mobilisé par pVCR94^{Sp}, vers une réceptrice contenant ou non SGI1^{Kn}. La présence de SGI1^{Kn} dans la réceptrice réduit le transfert de SGI1^{Cm} d'un facteur 100 environ (**figure 6A**). Cet îlot génomique pourrait donc encoder un système d'exclusion.

Les protéines de fonction inconnue encodées par SGI1 ont été alignées par BLAST (Altschul *et al.*, 1990) sur l'ensemble des génomes des Gammaprotéobactéries à la recherche d'homologues de protéines d'exclusion. S009 possède une faible similarité (44% d'identité sur 25 acides aminés, e-value de 0,61) avec une protéine TraS (accession NCBI : STX60751.1) retrouvée chez *Klebsiella pneumoniae*. Le gène *s009* est situé en aval de *traG* et encode une protéine de 96 acides aminés, prédite comme protéine de la membrane interne par PSORT (Gardy *et al.*, 2003). S009 possède ainsi plusieurs caractéristiques typiques d'un facteur d'exclusion.

Afin de confirmer l'implication de S009 dans l'exclusion de SGI1, une délétion de *s009* a été effectuée dans SGI1^{Kn}. Le transfert de SGI1^{Cm} mobilisé par pVCR94^{Sp} vers une réceptrice contenant SGI1^{Kn}Δ*s009* est environ dix fois plus efficace que lorsque la réceptrice contient SGI1^{Kn} (**figure 6A**). Cependant, la délétion de *s009* ne suffit pas à rétablir un taux de transfert similaire à celui observé vers une réceptrice vide.

Afin de savoir si la présence de S009 seul suffisait à induire une baisse du transfert de SGI1, *s009* a été cloné dans un plasmide pBAD30, sous le contrôle d'un promoteur inductible à l'arabinose et répressible au glucose (**figure 6B**). Le transfert de SGI1^{Kn} vers des cellules réceptrices contenant ce plasmide a été testé en présence d'arabinose et de glucose. On observe une baisse de transfert d'un facteur 10 environ en présence d'arabinose, c'est-à-dire lorsque *s009* est surexprimé.

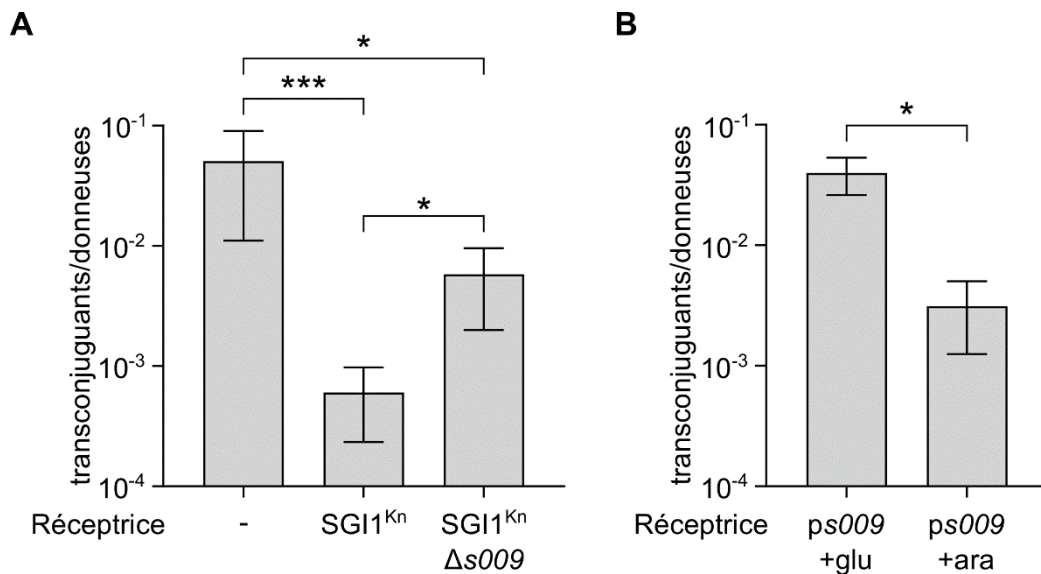


Figure 6. Rôle de S009 dans l'exclusion de SGI1

A : fréquence de transfert (en transconjugants/donneuses) de SGI1^{Cm} mobilisé par pVCR94^{Sp} vers des réceptrices ne contenant aucun îlot génomique ou qui portent SGI1^{Kn} ou SGI1^{Kn}Δs009. B : fréquence de transfert (en transconjugants/donneuses) de SGI1^{Kn} mobilisé par pVCR94^{Sp} vers une réceptrice contenant le plasmide ps009, en présence de glucose 0,2% (+glu) ou d'arabinose 0,2% (+ara).

Ces résultats confirment donc l'existence d'un système d'exclusion chez SGI1, impliquant S009 comme facteur d'exclusion. Ainsi, cette protéine est renommée EexG.

3.2.2. Exclusion de surface

3.2.2.1 *HDTM*

Nous avons observé précédemment que la délétion des gènes *vcrx085* à *vcrx141* avait un impact sur l'exclusion (2.2 Article : Entry Exclusion of Conjugative Plasmids of the IncA, IncC, and

Related Untyped Incompatibility Groups, page 36). De plus, la délétion de *eexC* n'est pas suffisante pour éliminer totalement l'exclusion exercée par pVCR94. Ces deux observations m'ont amené à émettre l'hypothèse que les plasmides IncC possèderaient au moins un autre gène d'exclusion, éventuellement responsable d'une exclusion de surface.

Afin d'identifier cet autre gène d'exclusion, Kévin Huguet, étudiant postdoctoral au sein de notre laboratoire, a effectué une mutagenèse par insertion de transposon (High-Density Transposons Mutagenesis, ou HDTM). Les étapes de cette expérience sont les suivantes :

- 1) Construction d'une librairie d'insertion de transposons Tn5 dans une souche BW25113Nx portant pVCR94^{Kn} Δ *eex::cat*. L'objectif est d'obtenir un million de clones, pour une densité d'insertion espérée d'environ une insertion toutes les cinq paires de bases. Si le transposon Tn5 s'insère à l'intérieur d'un gène, il interrompt la transcription du gène et de l'opéron à l'aide d'un terminateur fort. On emploie pVCR94^{Kn} Δ *eexC* parce qu'il est déjà connu que *eexC* est un facteur d'exclusion, et sa présence pourrait interférer avec la détection d'un deuxième système d'exclusion.
- 2) Sélection sur acide nalidixique (BW25113Nx), kanamycine (pVCR94^{Kn}), chloramphénicol (Δ *eex::cat*) et spectinomycine (transposon Tn5) afin de conserver uniquement les réceptrices qui n'ont pas perdu pVCR94^{Kn} et dans lesquelles le transposon Tn5 s'est intégré.
- 3) Transfert de pClo mobilisé par pVCR94^{Sp} vers ces réceptrices ayant reçu le transposon.
- 4) Sélection sur acide nalidixique (BW25113Nx), kanamycine (pVCR94^{Kn}), chloramphénicol (Δ *eex::cat*), spectinomycine (transposon Tn5) et ampicilline (pClo) pour conserver uniquement les réceptrices ayant reçu pClo. Les réceptrices dans lesquelles des gènes d'exclusion ont été interrompus devraient recevoir plus facilement

pClo, et donc être surreprésentées dans cette population. Ainsi, ces gènes devraient avoir une densité d'insertions plus importante à l'étape 4 qu'à l'étape 2.

- 5) Préparation de la banque d'HDTM, puis séquençage du locus d'insertion des transposons, pour les bactéries obtenues à l'étape 2 et à l'étape 4.

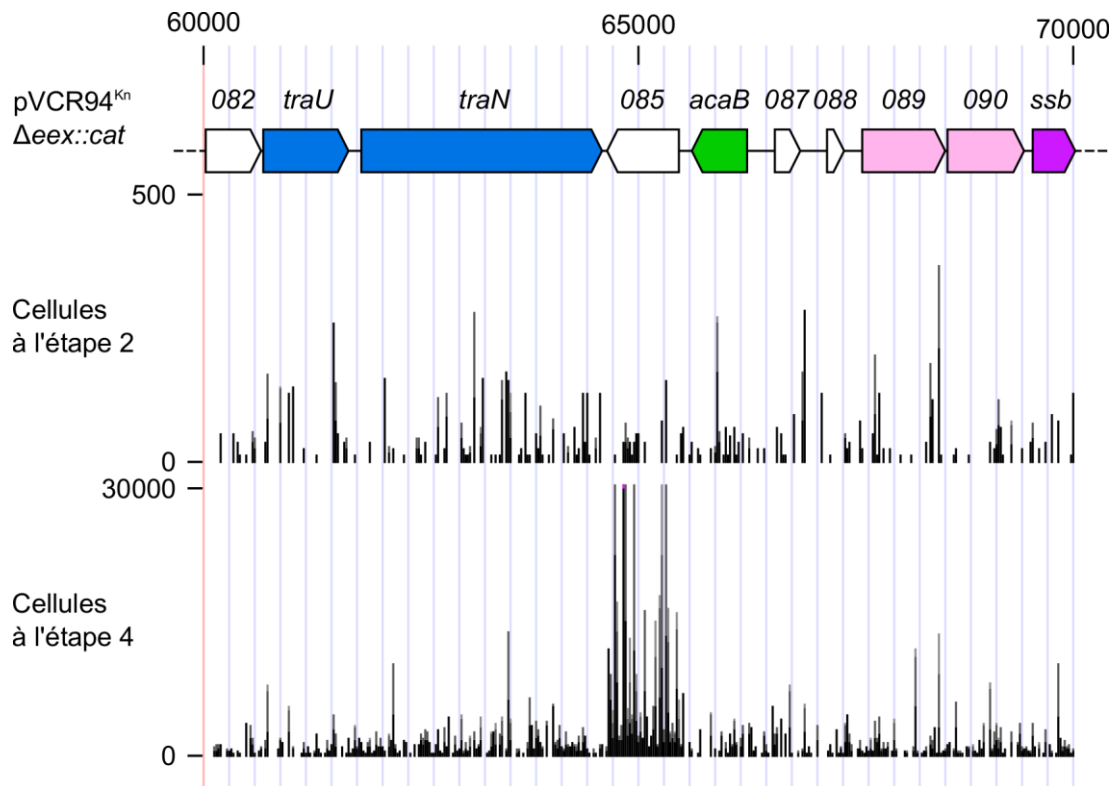


Figure 7. Identification par HDTM d'un potentiel déterminant de l'exclusion

Densité d'insertion de transposons dans pVCR94^{Kn}Δ*eex::cat* avant (étape 2) et après (étape 4) transfert de pClo. Seule la région de pVCR94^{Kn} située entre 60 kb et 70 kb est représentée

La densité d'insertions du transposon observée sur la totalité de pVCR94^{Kn} Δ *eex::cat* est présentée en **Figure S1** (page 94). Afin d'en améliorer la lisibilité, la **figure 7** ne présente que la portion de pVCR94 située entre 60 kb et 70 kb.

Le gène *vcrx085* présente une densité de transposons bien plus élevée à l'étape 4 qu'à l'étape 2. Sur le reste du plasmide, aucun autre gène ne présente une différence aussi visible (**Figure S1**, page 94). Ce gène est situé dans la délétion Δ D utilisée dans l'article, qui a un impact sur l'exclusion. Il semble donc être un bon candidat pour être un déterminant de l'exclusion.

3.2.2.2 Phylogénie

Pour savoir si Vcrx085 est encodé par d'autres plasmides et s'il est conservé parmi les plasmides IncC, j'ai effectué une analyse phylogénétique de sa séquence protéique.

Dans ce but, la séquence de Vcrx085 a été comparée aux génomes de toutes les Gammaprotéobactéries. Cette comparaison a été effectuée par l'algorithme tBLASTn du NCBI (Altschul *et al.*, 1990), qui permet de comparer une séquence protéique aux traductions obtenues avec tous les cadres de lectures des génomes étudiés. Ainsi, les résultats obtenus ne dépendent pas des annotations des protéines de ces génomes. Parmi toutes les séquences nucléotidiques obtenues, 317 présentaient au minimum 90% de couverture et 30% d'identité. Ces 317 séquences ont été traduites par Transeq (Madeira *et al.*, 2019), et regroupées en clusters de séquences identiques avec CDHit (Li et Godzik, 2006). Un seul représentant de chaque cluster est indiqué sur l'arbre phylogénétique ci-dessous, mais la liste complète des séquences utilisées est détaillée dans le **Tableau S1** (page 76).

Les représentants uniques des 317 protéines ont été alignées avec MUSCLE (Edgar, 2004), trimées avec TrimAL (Capella-Gutiérrez *et al.*, 2009), et l'arbre phylogénétique a été construit par MEGA (Kumar *et al.*, 2016) (**figure 8**).

On observe que Vcrx085 est très conservé parmi les plasmides IncC et IncA (pRA1, pIncAC-KP4898). Un homologue est également présent chez pAsa4c, pAhD4-1 et d'autres plasmides relativement proches des plasmides IncC, tels que pKHM-1 ou pXNC1.

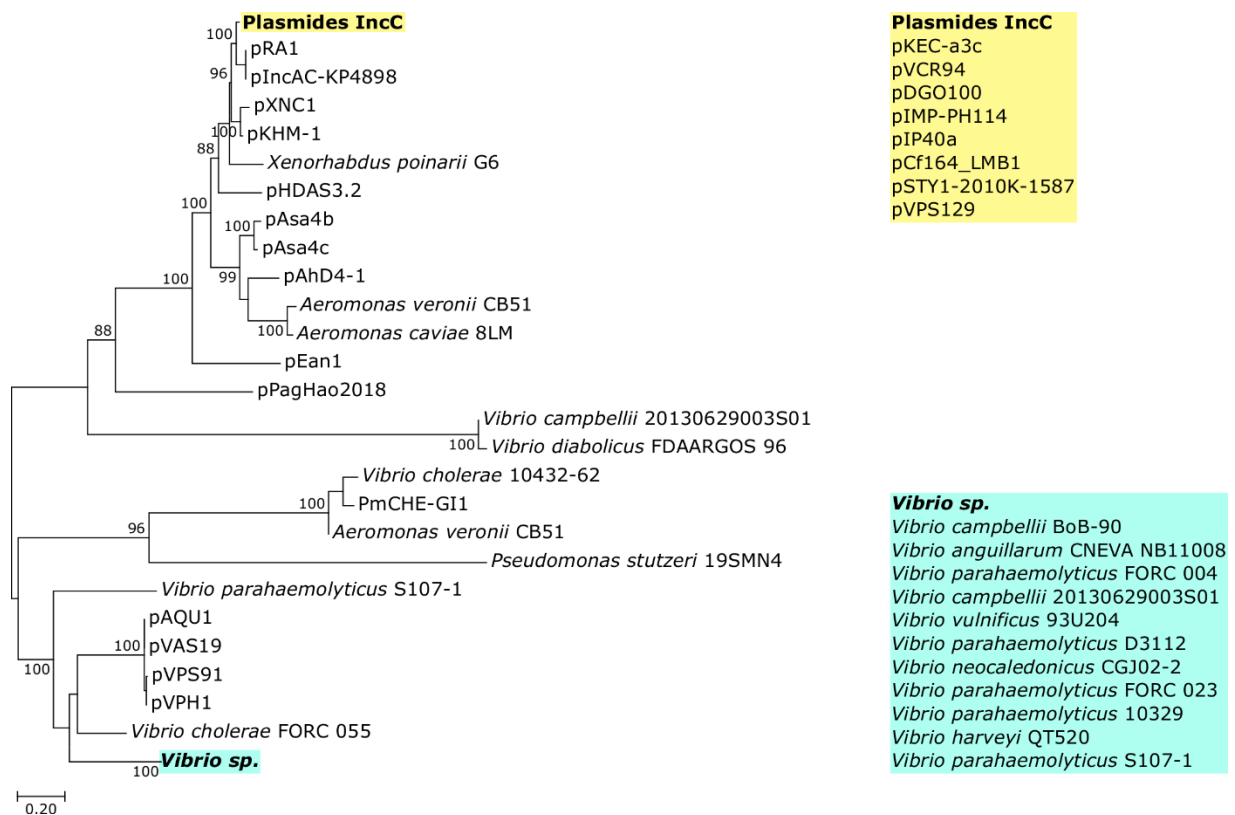


Figure 8. Analyse phylogénétique des ORFs homologues à vcrx085

Les bootstraps sont indiqués en pourcentages aux points de branchements seulement s'ils sont supérieurs à 80%. Un seul représentant de chaque cluster est montré dans cet arbre. Les plasmides IncC sont regroupés en une seule branche (en jaune), de même qu'un groupe de génomes du genre *Vibrio* (*Vibrio sp.*, en bleu).

On retrouve également un gène homologue à vcrx085 dans de nombreux génomes de *Vibrio* ainsi que chez *Pseudomonas* ou *Aeromonas*. Ces gènes sont présents dans des îlots génomiques

portés par ces bactéries, dont on ne sait pas s'ils sont mobilisables ou non par les plasmides IncC.

3.2.2.3 Prédiction de structure protéique

D'après l'analyse de sa séquence par PSORT-b (Gardy *et al.*, 2003), Vcrx085 serait non-cytoplasmique, et contiendrait un peptide signal. Selon SignalP-5.0 (Almagro Armenteros *et al.*, 2019), il s'agit d'un peptide signal de lipoprotéine transportée par le translocon Sec et clivée par la signal peptidase II, dont le site de clivage serait situé entre les acides aminés G18 et C19 (probabilité de 0,9984). La structure secondaire de Vcrx085 a été prédite par PSI-PRED (Jones, 1999) et est présentée dans la **figure 9**.

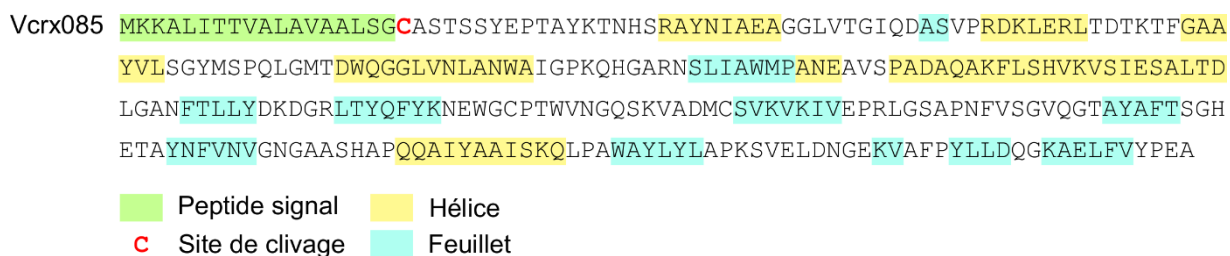


Figure 9. Prédiction de la structure de Vcrx085

Séquence de Vcrx085, et prédiction du peptide signal (en vert), du site de clivage (entre G18 et C19, en rouge), des hélices (en jaune) et des feuillets (en bleu).

3.2.2.4 Essais d'exclusion avec Vcrx085

Pour étudier l'impact de *vcrx085* sur l'exclusion des plasmides IncC, j'ai effectué le transfert de pClo, mobilisé par pVCR94^{Kn}, vers des réceptrices contenant pVCR94^{Sp}, pVCR94^{Sp}Δ*eexC*, pVCR94^{Sp}Δ*vcrx085* ou pVCR94^{Sp}Δ*eexC*Δ*vcrx085* (**figure 10A**). On observe qu'avec la délétion de *vcrx085* dans la réceptrice, on obtient un taux de transfert similaire à celui obtenu avec la délétion de *eexC*, c'est-à-dire plus élevé que vers une réceptrice contenant pVCR94^{Sp}

mais plus faible que vers une réceptrice ne contenant aucun plasmide. La double délétion $\Delta eexC\Delta vcrx085$, en revanche, rétablit un taux de transfert similaire à celui obtenu vers une réceptrice vide. Puisque pClo est compatible avec les plasmides IncC, on peut conclure que Vcrx085 est impliqué dans l'exclusion des plasmides IncC, et qu'il exerce une exclusion d'un niveau similaire à celle exercée par EexC.

Afin de vérifier que la présence de Vcrx085 dans la réceptrice est suffisante pour exercer cette exclusion, j'ai surexprimé *vcrx085* dans des réceptrices ne contenant pas pVCR94. Pour cela, j'ai cloné *vcrx085* dans un plasmide pBAD30, sous contrôle du promoteur P_{BAD}, inductible à l'arabinose. Ce plasmide est appelé p*vcrx085*. Le transfert de pVCR94^{Sp} vers des réceptrices contenant ou non p*vcrx085* a été effectué : le taux de transfert de pVCR94^{Sp} (en transconjugants/donneuse) diminue d'un facteur 10 environ lorsque *vcrx085* est surexprimé dans la réceptrice (**figure 10B**). Cependant, on observe également une baisse importante du nombre de réceptrices lors de la surexpression de ce gène (**figure 10C**). Il semble donc que la surexpression de *vcrx085* soit létale pour les cellules réceptrices, ou qu'elle ralentisse considérablement leur croissance.

Néanmoins, les résultats obtenus dans la **figure 10A** permettent de valider que *vcrx085* est impliqué dans l'exclusion des plasmides IncC. L'exclusion de surface étant généralement désignée par SFX dans la littérature (Thomas *et al.*, 2017), cette protéine est renommée SfxC.

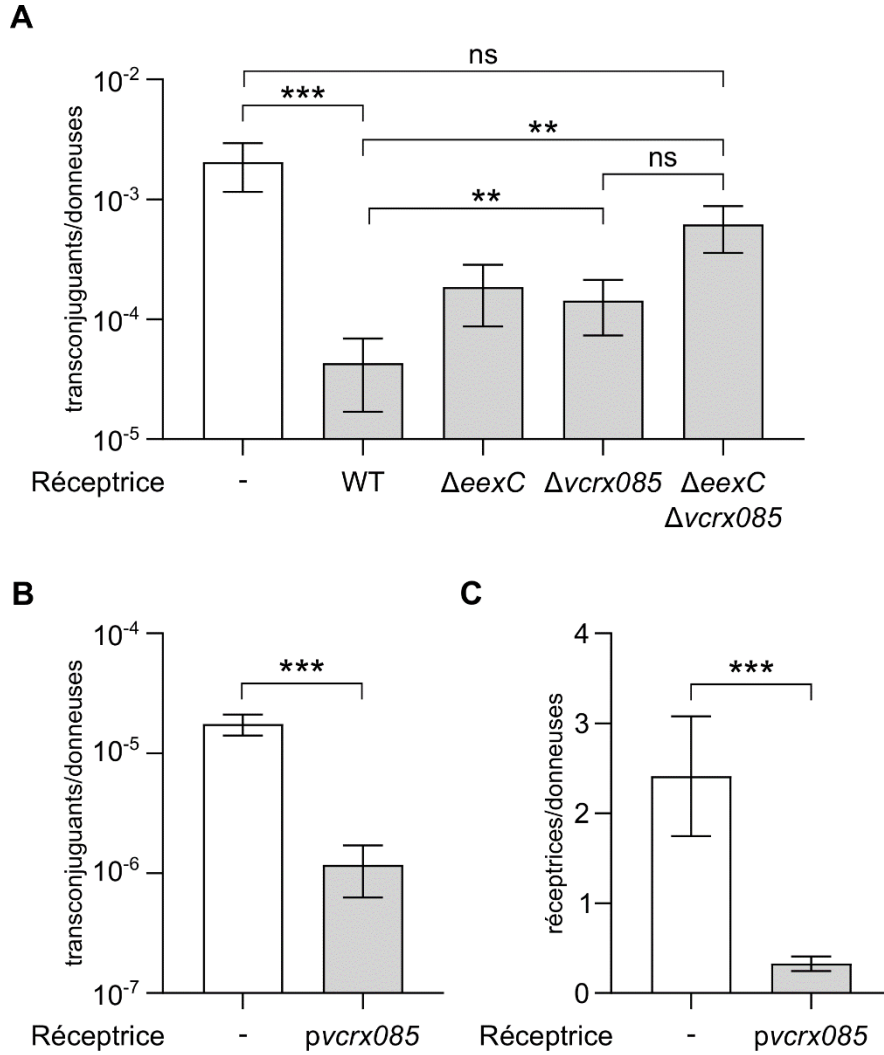


Figure 10. Rôle de *vcrx085* dans l'exclusion des plasmides IncC

A : taux de transfert (en transconjugants/donneuses) de pClo en présence de pVCR94^{Kn} vers une réceptrice vide ou contenant pVCR94^{Sp} (WT), pVCR94^{Sp}ΔeexC, pVCR94^{Sp}Δvcrx085 ou pVCR94^{Sp}ΔeexCΔvcrx085. B : taux de transfert (en transconjugants/donneuses) de pVCR94^{Sp} en absence ou en présence de pvcrx085 dans la réceptrice. C : ratio de réceptrices/donneuses en absence ou en présence de pvcrx085 dans la réceptrice.

CHAPITRE 4

DISCUSSION GÉNÉRALE ET CONCLUSION

Pendant cette maîtrise, j'ai avancé sur les points suivants :

- J'ai identifié les facteurs d'exclusion des plasmides IncC et IncA : EexC est responsable de l'exclusion d'entrée et SfxC de l'exclusion de surface.
- J'ai montré que TraG était la protéine du T4SS reconnue par EexC lors de l'exclusion d'entrée.
- J'ai déterminé qu'il existait au moins trois groupes d'exclusion d'entrée parmi les plasmides proches des plasmides IncC. On retrouve EexC et TraG_C chez les plasmides IncC et IncA, EexD et TraG_D chez pAhD4-1, ainsi que EexE et TraG_E chez pAsa4c et les plasmides apparentés. La reconnaissance entre TraG et les 20 acides aminés en C-terminal de Eex est responsable de la spécificité de ces groupes d'exclusion.
- J'ai mis en évidence que SGI1 contourne l'exclusion d'entrée des plasmides IncC en remplaçant TraG_C par TraG_S dans le pore de conjugaison.
- J'ai identifié un facteur d'exclusion d'entrée chez SGI1.

4.1 Exclusion d'entrée des plasmides IncC

Au cours de ma maîtrise, j'ai montré que EexC était le facteur d'exclusion d'entrée des plasmides IncC et IncA. Des homologues de cette protéine sont aussi retrouvés dans des plasmides non-typés, tels que pAhD4-1 ou pAsa4c. La reconnaissance de TraG par Eex permet

le blocage du transfert d'ADN lors de la conjugaison. Lors de sa mobilisation par un plasmide IncC, SGI1 remplace TraG_C par TraG_S, empêchant ainsi cette reconnaissance. De cette manière, SGI1 peut se disséminer dans une population de bactéries qui portent des plasmides IncC ou IncA.

Des différences de séquences entre les Eex et TraG encodés par les plasmides qui les portent définissent trois groupes d'exclusion : les plasmides IncC et IncA (EexC/TraG_C) s'excluent mutuellement, mais n'excluent pas d'autres plasmides tels que pAhD4-1 (EexD/TraG_D) ou pAsa4c (EexE/TraG_E).

Ce système d'exclusion d'entrée est similaire à ceux des plasmides IncF (Audette *et al.*, 2007), des plasmides IncI (Sakuma *et al.*, 2013) et des ICE SXT/R391 (Marrero et Waldor, 2005). En effet, dans tous ces systèmes, le facteur d'exclusion est une protéine ancrée dans la membrane interne de la cellule réceptrice, et interagit avec un homologue de TraG (nommé TraG chez le plasmide F et les ICE SXT/R391, et TraY chez les plasmides IncI). L'exclusion ne nécessite pas la présence d'autres protéines encodées par le plasmide dans la réceptrice. Plusieurs groupes d'exclusion existent dans chacun de ces systèmes, dus à la reconnaissance de TraG ou TraY par le facteur d'exclusion.

Chez les plasmides IncC, j'ai identifié les 20 acides aminés en C-terminal de Eex comme étant responsables de la spécificité des groupes d'exclusion. Selon des prédictions bio-informatiques, cette portion de la protéine serait située dans le cytoplasme de la cellule réceptrice (2.2 Article : Entry Exclusion of Conjugative Plasmids of the IncA, IncC, and Related Untyped Incompatibility Groups, page 36). Marrero et Waldor (2007b) ont montré que chez les ICE SXT/R391, les portions de EexS et EexR permettant la reconnaissance de TraG sont situées dans le cytoplasme de la réceptrice. De plus, seuls trois acides aminés de TraG_S et de TraG_R sont responsables de la spécificité des groupes d'exclusion, et ces résidus sont situés dans le cytoplasme de la donneuse. À ce jour, aucune étude n'a pu montrer de quelle manière ces deux protéines peuvent interagir bien qu'elles soient séparées par les parois des bactéries donneuses

et réceptrice. Chez le plasmide F, les tentatives d'immunoprécipitation, de cross-linking ou de double-hybride bactérien pour mettre en évidence cette interaction ont été infructueuses (Audette *et al.*, 2007).

Il est possible qu'une portion ou la totalité de TraG soit transloquée vers la cellule réceptrice, par autotranslocation ou à travers le T4SS, par exemple. Dans ce dernier cas, l'interaction de Eex avec TraG pourrait provoquer l'arrêt de cette translocation par encombrement du canal central du T4SS et donc le blocage du pore de conjugaison. Une autre hypothèse, développée par Joeli Marrero dans sa thèse de doctorat (Marrero, 2007), suppose que le transfert conjugatif nécessite la translocation de TraG, la reconnaissance par celui-ci d'un récepteur dans la membrane interne de la cellule réceptrice, et la transduction d'un signal jusqu'à la cellule donneuse, permettant l'initiation du transfert. Dans ce contexte, Eex pourrait bloquer l'interaction de TraG avec son récepteur, et empêcher l'envoi du signal d'initiation de la conjugaison.

4.2 Exclusion d'entrée de SGI1

À l'instar des plasmides IncC, SGI1 possède un système d'exclusion d'entrée. En effet, il encode la protéine EexG, qui, lorsqu'elle est présente dans la cellule réceptrice, empêche l'entrée d'une autre copie de SGI1. Cette exclusion fonctionne uniquement si TraG_S est présent dans le T4SS. L'implication de TraG_S dans ce mécanisme suggère que ce système d'exclusion pourrait fonctionner d'une manière similaire à celui des plasmides IncC. Cependant, il n'y a pas d'homologie entre EexC et EexG.

Il s'agit de la première observation d'un phénomène d'exclusion causé par îlot génomique mobilisable. On peut supposer que de tels système sont rares, car le processus d'exclusion nécessite la reconnaissance par la réceptrice d'une protéine encodée par l'îlot, avant le transfert de l'ADN de celui-ci. La modification du T4SS des plasmides IncC par SGI1 a probablement grandement facilité l'apparition de ce système d'exclusion. À ce jour, SGI1 et ses variants sont

les seuls MGI connus à être capable de modifier ainsi le pore de conjugaison du plasmide qui les mobilise (Carraro *et al.*, 2017a).

Un système d'exclusion pourrait exister chez les plasmides ColE1 qui sont mobilisables par une grande variété de plasmides conjuguatifs (Finnegan et Sherratt, 1982). Cependant, d'après Garcillán-Barcia et de la Cruz (2008), les résultats obtenus jusqu'à présent ne sont pas suffisants pour s'assurer qu'il s'agit bel et bien d'exclusion et pas d'incompatibilité. En dehors des plasmides ColE1, aucun autre système d'exclusion n'a été identifié chez un élément mobilisable.

Même en l'absence du facteur d'exclusion EexG, la présence de SGI1 Δ eexG dans la réceptrice défavorise le transfert d'une autre copie de SGI1 vers cette bactérie. Ceci pourrait être dû à l'absence d'un site *attB* dans le chromosome à la suite de l'intégration d'une première copie de SGI1. En effet, cette première intégration dans le site *attB* mène à la formation des sites *attL* et *attR*. Ces sites n'étant pas identiques à *attB*, cela pourrait compliquer l'intégration d'une nouvelle copie de SGI1. Cependant, SGI1 est capable de s'insérer dans un deuxième site *attB*, situé entre les gènes *sodB* et *purR*, ainsi que de s'insérer en tandem, c'est-à-dire avec plusieurs copies dans le même site (Doublet *et al.*, 2008). Il est donc probable que la baisse de transfert observée en présence de SGI1 Δ eexG dans la réceptrice ne soit pas due à une difficulté lors de son intégration, mais plutôt à la présence d'un deuxième facteur d'exclusion encodé par SGI1.

4.3 Exclusion de surface de pVCR94

Il existe un deuxième système d'exclusion chez les plasmides IncC, qui implique SfxC. Tout comme TraT, la protéine d'exclusion de surface du plasmide F, SfxC possède un signal peptide qui pourrait permettre son ancrage lipidique dans la membrane externe. Cependant, je n'ai pas encore pu vérifier que SfxC était bien une protéine de la membrane externe. Pour s'assurer que l'exclusion médiée par SfxC est une exclusion de surface, il faudrait donc vérifier sa localisation.

Il faudrait également vérifier que sa présence dans la cellule réceptrice empêche l'établissement ou le maintien de paires de conjugaison stables, comme c'est le cas avec TraT.

Il serait aussi intéressant de savoir avec quelle(s) protéine(s) de la cellule donneuse SfxC interagit. Chez le plasmide F, plusieurs hypothèses ont été avancées. Premièrement, TraT aurait pu interagir avec la piline : cette hypothèse a été réfutée par Anthony *et al.* (1994). Ces mêmes auteurs ont proposé l'existence d'une interaction entre TraT et une adhésine qui serait située à l'extrémité du pilus, mais n'ont pas vérifié cette hypothèse (Anthony *et al.*, 1994). Klimke et Frost (1998) ont montré que chez la donneuse, TraN est responsable de la reconnaissance de la cellule réceptrice et de la stabilisation de la paire de conjugaison, en interagissant avec les protéines OmpA à la surface de la réceptrice. Ils ont émis l'hypothèse que l'interaction TraN-OmpA serait déstabilisée par TraT. Cependant, les résultats obtenus dans cette étude n'ont montré aucune implication de TraN dans l'exclusion de surface. Néanmoins, cette conclusion se base sur des résultats obtenus uniquement avec TraT_F, et pas avec les protéines d'exclusion des autres groupes d'exclusion des plasmides IncF, tels que R100-1. En ce qui concerne les plasmides IncC, il me semble toutefois intéressant de vérifier le rôle de TraN dans l'exclusion médiée par SfxC. En effet, de manière générale, les gènes d'exclusion sont souvent localisés à proximité du gène encodant leur partenaire d'interaction (Garcillán-Barcia et de la Cruz, 2008). Or, *sfxC* est situé directement en aval de *traN*. De plus, SGI1 remplace TraN dans le T4SS des plasmides IncC sans que cela améliore son transfert vers une réceptrice ne contenant pas de plasmide IncC (Carraro *et al.*, 2017a). On peut donc imaginer que TraNs apporte un avantage à SGI1 en lui permettant de s'évader du deuxième système d'exclusion des plasmides IncC.

Les réceptrices dans lesquelles *sfxC* est surexprimé présentent une surmortalité ou un important retard de croissance. Cela peut être dû à la présence en trop grande quantité de SfxC dans la membrane externe, qui entraînerait une déstabilisation de cette dernière. À ma connaissance, ceci n'a pas été observé chez le plasmide F, mais dans des conditions d'expressions normales, on retrouve 20 000 à 80 000 copies de TraT par cellule (Achtman *et al.*, 1977). TraT est également impliqué dans la virulence des bactéries qui le portent en leur permettant de mieux

résister au système immunitaire de leur hôte. Garcillán-Barcia et de la Cruz (2008) émettent l'hypothèse que la fonction première de TraT soit de modifier la surface de la cellule pour améliorer leur virulence. On peut donc supposer qu'en trop grande quantité, une telle protéine provoque une déstabilisation de la membrane externe.

Des homologues de SfxC sont retrouvés dans des îlots génomiques chez diverses espèces de *Vibrio*, *Aeromonas* ou *Pseudomonas*. La mobilisation des îlots présentés dans la **figure 8** par les plasmides IncC n'a pas été testée. Cependant, un homologue plus distant, qui n'est pas représenté dans l'arbre phylogénétique, est retrouvé chez MGII*ch*Hai6, un îlot mobilisé par les plasmides IncC. Étant donné les fortes différences entre ces protéines et SfxC, il est peu probable que ces îlots puissent exclure les plasmides IncC. En revanche, il se pourrait que, comme TraT, SfxC possède une autre fonction bénéfique pour la bactérie qui le porte, ce qui expliquerait qu'il soit retrouvé dans de nombreux éléments.

4.4 Conclusion et perspectives

Ainsi, les plasmides IncC encodent deux facteurs d'exclusion, EexC et SfxC. Ces protéines sont retrouvées également chez les plasmides IncA et d'autres plasmides non-typés, tels que pAsa4c ou pAhD4-1. L'étude de cette exclusion permet de mieux connaître et comprendre ces éléments, qui sont des facteurs majeurs de la dissémination de multirésistances aux antibiotiques.

Cette étude a également permis d'ajouter une nouvelle dimension à la relation déjà complexe entre SGI1 et les plasmides IncC. L'îlot ne se contente pas d'utiliser le système conjugatif de ces plasmides, mais le modifie à son avantage, et au détriment des derniers. En contournant au moins l'un des systèmes d'exclusion de ces plasmides, SGI1 peut se disséminer efficacement dans une population contenant des plasmides IncC. L'îlot est également capable d'améliorer sa propre stabilité en déstabilisant ces plasmides (Huguet *et al.*, 2020). Ainsi, bien que les plasmides IncC soient nécessaires à la dissémination de SGI1, ces deux éléments sont rarement isolés dans une même souche bactérienne.

L'exclusion est un moyen efficace de réduire le transfert conjugatif. En comprenant mieux ce mécanisme, il serait peut-être possible de mettre au point des composés anti-conjugaison. Ces composés pourraient éventuellement être utilisés en parallèle des antibiotiques pour limiter la dissémination de résistances chez un patient sous traitement antibiotique.

Pour comprendre plus en profondeur les mécanismes d'exclusion d'entrée et de surface, de nouvelles études seront nécessaires. Des études génétiques pourraient permettre d'identifier la cible de SfxC chez les donneuses. Des études structurales pourraient permettre de mieux comprendre l'interaction entre les différents facteurs impliqués dans l'exclusion. La compréhension de l'interaction entre Eex et TraG, notamment, pourrait aider à mieux comprendre le mécanisme de conjugaison en lui-même.

ANNEXES

Tableau S1. Clusters de protéines Vcrx085 identiques

Les protéines homologues à Vcrx085 ont été identifiées par une recherche tBLASTn sur la base de données des Gammaprotéobactéries (identité > 15%, longueur > 90%). Les clusters de protéines identiques ont été formés par CD-Hit (seuil d'identité de 1).

Cluster	Plasmide	Groupe Inc	Accession	Espèce	Souche	Longueur (aa)
pVCR94	pVCR94	IncC	CP033514.1	<i>Vibrio cholerae</i>	E4	286
	pEA409TEM24	IncC	MG764534.2	<i>Klebsiella aerogenes</i>	EA409	286
	pMG252	IncC	MK638972.1	<i>Escherichia coli</i>	J53	286
	p23-C-23	IncC	CP038466.1	<i>Aeromonas hydrophila</i>	23-C-23	286
	pWCX23-2	IncC	CP038464.1	<i>Aeromonas hydrophila</i>	WCX23	286
	pEco190	IncC	CP020524.1	<i>Escherichia coli</i>	190	286
	pC-Ec2-KPC	IncC	MK439959.1	<i>Escherichia coli</i>	Ec-2Lar	286
	p526316-KPC	IncC/IncQ1	MH909327.1	<i>Klebsiella pneumoniae</i>	526316	286
	pC-Ec20-KPC	IncC	MK388209.1	<i>Escherichia coli</i>	Ec20-Lar	286
	pBA34918-IncAC2	IncC	CP036192.1	<i>Klebsiella pneumoniae</i>	BA34918	286
	pECL17464	IncC	MH995508.1	<i>Enterobacter cloacae</i>	ECL17464	286
	pCFR17394	IncC	MH995506.1	<i>Citrobacter freundii</i>	CFR17394	286
	pSenAgo44	IncC	MK191845.1	<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar Agona	44	286
	p1_GR_13-IncAC2	IncC	CP027043.1	<i>Klebsiella pneumoniae</i>	1_GR_13	286
	p_2_GR_12- IncAC2	IncC	CP027055.1	<i>Klebsiella pneumoniae</i>	2_GR_12	286
	p16_GR_13- IncAC2	IncC	CP027038.1	<i>Klebsiella pneumoniae</i>	16_GR_13	286
	pKpn16_1	IncC	LT968702.1	<i>Klebsiella pneumoniae</i>	16_1	286
	pNDM-2262	IncC	MH892479.1	<i>Citrobacter freundii</i>	2262	286
	pNDM-2248	IncC	MH844629.1	<i>Escherichia coli</i>	2248	286
	p2016K-0796	IncC	MH760469.1	<i>Salmonella enterica</i>	2016K-0796	286
	p22429	IncC/IncFIA /IncFIB/Inc FII/IncX1	CP032397.1	<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar Dublin	CVM_22429	286

Cluster	Plasmide	Groupe Inc	Accession	Espèce	Souche	Longueur (aa)
	p34981_1	IncC	CP032391.1	<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar Dublin	CVM_34981	286
	pTB221	IncC	CP032238.1	<i>Escherichia coli</i>	ECCWS199	286
	pAR_0127-1	IncC	CP032192.1	<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar Senftenberg	AR_0127	286
	pAR_0076-1	IncC	CP032168.1	<i>Klebsiella pneumoniae</i>	AR_0076	286
	pIA/C-KLUC	IncC	MH476540.1	<i>Klebsiella pneumoniae</i>	KP1276	286
	pCFJY-17	IncC	MH763829.1	<i>Citrobacter freundii</i>	JY-17	286
	pIncAC2_L111	IncC	CP030132.1	<i>Klebsiella pneumoniae</i>	160111	286
	pCCU101-1	IncC	CP018457.1	<i>Shewanella algae</i>	CCU101	286
	pC1394	IncC	MH457126.1	<i>Vibrio alginolyticus</i>	Vb1394	286
	pIncAC2-1502318	IncC	CP031610.1	<i>Escherichia coli</i>	S10	286
	pIncAC2-1502320	IncC	CP031584.1	<i>Klebsiella pneumoniae</i>	S12	286
	pIncAC2-1301491	IncC	CP031570.1	<i>Enterobacter hormaechei</i>	S13	286
	pIncAC2-1502264	IncC	CP031576.1	<i>Enterobacter hormaechei</i>	S6	286
	pIncAC2-1502262	IncC	CP031573.1	<i>Enterobacter hormaechei</i>	S5	286
	pFDRGOS_87-1	IncC	CP031512.1	<i>Providencia stuartii</i>	FDRGOS_87	286
	pBA4656-IncAC2	IncC	CP035908.1	<i>Klebsiella pneumoniae</i>	BA4656	286
	pSNE2-2012K-0663	IncC	CP025245.1	<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar Newport	CDC_2012K-0663	286
	pSNE3-1928	IncC	CP025240.1	<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar Newport	USDA-ARS-USMARC-1928	286

Cluster	Plasmide	Groupe Inc	Accession	Espèce	Souche	Longueur (aa)
	pSNE1-1924	IncC	CP025231.1	<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar Newport	USDA-ARS-USMARC-1924	286
	pCRE7-NDM	IncC	MK101346.1	<i>Citrobacter freundii</i>	CRE3	286
	p427113-KPC	IncC/IncQ1	MH917285.1	<i>Klebsiella pneumoniae</i>	427113	286
	pNDM-M21015	IncC	MK041212.1	<i>Citrobacter amalonaticus</i>	M21015	286
	pWCX23_1	IncC	CP028419.1	<i>Aeromonas hydrophila</i>	WCX23	286
	pGD31-NDM	IncC	CP031297.1	<i>Escherichia coli</i>	EC17GD31	286
	pCq9-3	IncC	CP031549.1	<i>Escherichia coli</i>	cq9	286
	pKP30835-2	IncC	CP027695.1	<i>Klebsiella pneumoniae</i>	KP30835	286
	pIMP69_000369	IncC	CP031122.1	<i>Providencia sp.</i>	WCHPr000369	286
	pAMSC1	IncC	CP031106.1	<i>Escherichia coli</i>	AMSCJX02	286
	pSA20025921.1	IncC	CP030215.1	<i>Salmonella enterica</i>	SA20025921	286
	p3-20710	IncC	CP030077.1	<i>Enterobacter hormaechei</i>	20710	286
	pAR_0082-1	IncC	CP029737.1	<i>Providencia rettgeri</i>	AR_0082	286
	pTVGHCRE225-2	IncC	CP023724.1	<i>Klebsiella pneumoniae</i>	TVGHCRE225	286
	pPmi70746_1	IncC	CP023274.1	<i>Proteus mirabilis</i>	CCUG_70746	286
	pGSU10-3-1	IncC	AP018672.1	<i>Klebsiella pneumoniae</i>	GSU10-3	286
	pNDM-QD51	IncC	MH263652.1	<i>Providencia rettgeri</i>	QD51	286
	pKPC_CAV2018-435	Untyped	CP029431.1	<i>Klebsiella quasipneumoniae</i>	CAV2018	286
	pKPC_CAV2013	Untyped	CP029436.1	<i>Klebsiella quasipneumoniae</i>	CAV2013	286
	pKPC_CAV1947-412	Untyped	CP029442.1	<i>Klebsiella quasipneumoniae</i>	CAV1947	286

Cluster	Plasmide	Groupe Inc	Accession	Espèce	Souche	Longueur (aa)
	pSL131_IncA/C-IncX3	IncC/IncX3	MH105050.1	<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar Lomita	SL131	286
	pVb1796	IncC	MH113855.1	<i>Vibrio alginolyticus</i>	Vb1796	286
	pHNAH8I-NDM	IncC	MH061195.1	<i>Klebsiella pneumoniae</i>	AHM7C8I	286
	pNDM-TAEC1	IncC	MH001166.1	<i>Escherichia coli</i>	TAEC1	286
	pAR435-5	IncC	CP029118.1	<i>Escherichia coli</i>	AR435	286
	pAR_0079-5	IncC	CP028996.1	<i>Klebsiella pneumoniae</i>	AR_0079	286
	pSMO-2010K-0257	IncC	CP020913.1	<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar Montevideo	CDC_2010K-0257	286
	pSE12-01738-2	IncC	CP027679.1	<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar Corvallis	12-01738	286
	pCMY2_085072	IncC	CP028804.1	<i>Klebsiella pneumoniae</i>	WCHKP7E2	286
	pEI-MS-17-156-1	IncC	CP028814.1	<i>Edwardsiella ictaluri</i>	MS-17-156	286
	pNDM-185	IncC	MH011352.1	<i>Klebsiella pneumoniae</i>	185	286
	pEsco-36073cz	IncC	MG252895.1	<i>Escherichia coli</i>	Esco-36073cz	286
	pAMA566	IncC	MG450360.1	<i>Escherichia coli</i>	AMA566	286
	pGMI17-002_1	IncC	CP028170.1	<i>Salmonella enterica</i>	CFSAN064034	286
	pRCS53TR724_p_1	IncC	LT985260.1	<i>Escherichia coli</i>	724	286
	pRCS54TR726_p_1	IncC	LT985258.1	<i>Escherichia coli</i>	726	286
	pRCS48TR660_p_1	IncC	LT985253.1	<i>Escherichia coli</i>	660	286
	pRCS38_p_1	IncC	LT985250.1	<i>Escherichia coli</i>	170	286
	pRCS46_p_1	IncC	LT985249.1	<i>Escherichia coli</i>	195	286
	pRCS41_p_1	IncC	LT985243.1	<i>Escherichia coli</i>	722	286

Cluster	Plasmide	Groupe Inc	Accession	Espèce	Souche	Longueur (aa)
	pRCS28TR552_p_1	IncC	LT985228.1	<i>Escherichia coli</i>	552	286
	pRCS2TR89_p_1	IncC	LT985225.1	<i>Escherichia coli</i>	89	286
	pRCS30_p_1	IncC	LT985224.1	<i>Escherichia coli</i>	513	286
	pRCS24TR548_p_1	IncC	LT985222.1	<i>Escherichia coli</i>	548	286
	pRCS1TR83_p_1	IncC	LT985220.1	<i>Escherichia coli</i>	83	286
	p205880-AC2	IncC	MF344573.1	<i>Klebsiella pneumoniae</i>	N201205880	286
	pCFR-0b27	IncC	CP026233.1	<i>Citrobacter freundii</i> complex sp.	CFNIH4	286
	pECO-c85f	IncC	CP026405.1	<i>Escherichia coli</i>	ECONIH4	286
	pKOR-01e8	IncC	CP026281.1	<i>Klebsiella oxytoca</i>	KONIH2	286
	pECO-dc1b	IncC	CP026207.1	<i>Escherichia coli</i>	ECONIH5	286
	pKPC-8e09	IncC	CP026225.1	<i>Aeromonas sp.</i>	ASNIH3	286
	pKKp4-VIM	IncC	MF582638.1	<i>Klebsiella pneumoniae</i>	KKp4	286
	p1540-2	IncC	CP019053.1	<i>Escherichia coli</i>	CRE1540	286
	pNR5632_p1	IncC	CP025144.1	<i>Klebsiella pneumoniae</i>	NR5632	286
	pKP1766_p1	IncC	CP025147.1	<i>Klebsiella pneumoniae</i>	KP1766	286
	pKP1768_p1	IncC	CP025141.1	<i>Klebsiella pneumoniae</i>	KP1768	286
	pINF274_p1	IncC	CP024571.1	<i>Klebsiella pneumoniae</i>	INF274	286
	pINF278_p1	IncC	CP024564.1	<i>Klebsiella pneumoniae</i>	INF278	286
	pINF164_p1	IncC	CP024557.1	<i>Klebsiella pneumoniae</i>	INF164	286
	pINF163_p1	IncC	CP024550.1	<i>Klebsiella pneumoniae</i>	INF163	286
	pKSB1_9D_p1	IncC	CP024536.1	<i>Klebsiella pneumoniae</i>	KSB1_9D	286
	pINF157_p1	IncC	CP024529.1	<i>Klebsiella pneumoniae</i>	INF157	286
	pINF158_p1	IncC	CP024522.1	<i>Klebsiella pneumoniae</i>	INF158	286
	pKSB1_5D_p1	IncC	CP024192.1	<i>Klebsiella pneumoniae</i>	KSB1_5D	286
	pVb0499	IncC	MF627445.1	<i>Vibrio parahaemolyticus</i>	Vb0499	286

Cluster	Plasmide	Groupe Inc	Accession	Espèce	Souche	Longueur (aa)
	pVb0267	IncC	MF627444.1	<i>Vibrio parahaemolyticus</i>	Vb0267	286
	pFDRGOS_447	IncC	CP023952.1	<i>Klebsiella pneumoniae</i>	FDRGOS_447	286
	pFDRGOS_446	IncC	CP023948.1	<i>Klebsiella pneumoniae</i>	FDRGOS_446	286
	pT212	IncC	CP017084.1	<i>Proteus mirabilis</i>	T21	286
	pP0855	IncC	LC225353.1	<i>Photobacterium damsela</i> subsp. <i>piscicida</i>		286
	pT1	IncC	KX147633.1	<i>Citrobacter freundii</i>	AMA332	286
	pVb0506	IncC	MF497432.1	<i>Vibrio alginolyticus</i>	Vb0506	286
	p112298-tetA	IncC	KY986974.1	<i>Citrobacter freundii</i>	112298	286
	p80-2002_p2	IncC	LT904892.1	<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar Typhi	80-2002	286
	p16Pre36-NDM	Untyped	KX832927.1	<i>Providencia rettgeri</i>	16pre36	286
	pSHE-CTX-M	IncC	CP022359.1	<i>Shewanella bicestrii</i>	JAB-1	286
	pEc19	IncC	KY887591.1	<i>Escherichia coli</i>	Ec19	286
	pKP64216b	IncC	MF150123.1	<i>Klebsiella pneumoniae</i>	A64216	286
	pKP64477c	IncC	MF150121.1	<i>Klebsiella pneumoniae</i>	A64477	286
	pPM64421a	IncC	MF150118.1	<i>Proteus mirabilis</i>	A64421	286
	pEc158	IncC	KY887596.1	<i>Escherichia coli</i>	Ec158	286
	pEc78	IncC	KY887595.1	<i>Escherichia coli</i>	Ec78	286
	pKp55	IncC	KY887594.1	<i>Klebsiella pneumoniae</i>	Kp55	286
	pCf53	IncC	KY887593.1	<i>Citrobacter freundii</i>	Cf53	286
	pEc9	IncC	KY887590.1	<i>Escherichia coli</i>	Ec9	286
	p1605752AC2	IncC	CP022126.1	<i>Klebsiella pneumoniae</i>	DHQP1605752_NV	286
	pAR_0148_tig00000169_pilon	IncC	CP021952.1	<i>Klebsiella pneumoniae</i>	AR_0148	286
	Chromosomique	NA	CP021960.1	<i>Klebsiella pneumoniae</i>	AR_0139	286

Cluster	Plasmide	Groupe Inc	Accession	Espèce	Souche	Longueur (aa)
	pAR_0055_unitig_2	IncC	CP021936.1	<i>Escherichia coli</i>	AR_0055	286
	pKpnAR0107	IncC	CP021956.1	<i>Klebsiella pneumoniae</i>	AR_0107	286
	pAR_0120_tig00000516_pilon	IncC	CP021835.1	<i>Klebsiella pneumoniae</i>	AR_0120	286
	pAR_0156_unitig_1	IncC	CP021853.1	<i>Proteus mirabilis</i>	AR_0156	286
	pAR_0143_tig00000001	IncC	CP021709.1	<i>Klebsiella pneumoniae</i>	AR_0143	286
	pAR_0155_tig00000123	IncC	CP021695.1	<i>Proteus mirabilis</i>	AR_0155	286
	pAR_0159_tig00000137	IncC	CP021551.1	<i>Proteus mirabilis</i>	AR_0159	286
	pM216_AC2	IncC	AP018145.1	<i>Escherichia coli</i>	M216_1	286
	pM214_AC2	IncC	AP018143.1	<i>Escherichia coli</i>	M214_1	286
	pENCL58_01	IncC	LT882699.1	<i>Enterobacter cloacae</i>	ENCL58	286
	pKLPN57_01	IncC	LT882698.1	<i>Klebsiella pneumoniae</i>	KLPN57	286
	pVC211	IncC	KY399978.1	<i>Vibrio cholerae O139</i>	ICDC-211	286
	Chromosomique	NA	CP020841.1	<i>Klebsiella pneumoniae</i>	KPN1482	286
	pR222	IncC	KX869741.1	<i>Enterobacter cloacae</i>	20130723	286
	pB557-NDM	IncC	KX786648.1	<i>Enterobacter cloacae</i>	B557	286
	pVVS1-per1	IncC	KY014466.1	<i>Vibrio vulnificus</i>		286
	pVPS114	IncC	KY014465.1	<i>Vibrio parahaemolyticus</i>		286
	pAR_0069_unitig_2	IncC	CP020056.1	<i>Escherichia coli</i>	AR_0069	286
	pECAZ155_KPC	IncC	CP019001.1	<i>Escherichia coli</i>	Ecol_AZ155	286
	pEC316_KPC	IncC	CP018956.1	<i>Escherichia coli</i>	Ecol_316	286
	pKP38731_p1	IncC	CP014295.1	<i>Klebsiella pneumoniae</i>	KP38731	286
	pRJ119-NDM1	IncC	KX636095.1	<i>Klebsiella pneumoniae</i>	RJ119	286

Cluster	Plasmide	Groupe Inc	Accession	Espèce	Souche	Longueur (aa)
	pPm14C18	IncC	KU605240.1	<i>Proteus mirabilis</i>	Pm14C18	286
	pCR14_2	IncC	CP015394.1	<i>Klebsiella pneumoniae</i>	CR14	286
	pAR_0049_unitig_1	IncC	CP018817.1	<i>Klebsiella pneumoniae</i>	AR_0049	286
	pECY53	IncC	KT997783.1	<i>Escherichia coli</i>	Y5	286
	pCFSAN003890	IncC	CP016013.1	<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar Newport	CFSAN003890	286
	pPS1	IncC	CP017055.1	<i>Providencia stuartii</i>	BE2467 CAUTI	286
	pK-109-R-1	IncC	KX029331.1	<i>Klebsiella pneumoniae</i>	K-109-R	286
	Chromosomique	NA	CP017671.1	<i>Providencia rettgeri</i>	RB151	286
	pMS6198A	IncC	CP015835.1	<i>Escherichia coli</i>	MS6198	286
	pSL157_p1	IncC	CP017058.1	<i>Citrobacter freundii</i>	SL151	286
	pB2-1	IncC	KX458222.1	<i>Klebsiella pneumoniae</i>	B2	286
	pAC5331	Untyped	KX064437.1	<i>Citrobacter freundii</i>	05K0870	286
	pR16a	IncC	KX156773.1	<i>Escherichia coli</i>	K-12	286
	pKP1-NDM-1	IncC	KF992018.2	<i>Klebsiella pneumoniae</i>	KP1	286
	pEC2-NDM-3	IncC	KC999035.4	<i>Escherichia coli</i>	EC2	286
	pASP-a58	IncC	CP014775.1	<i>Aeromonas veronii</i>	AVNIH1	286
	pSTY1-1896	IncC	CP014978.1	<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar Typhimurium	USDA-ARS-USMARC-1896	286
	pSAN1-1736	IncC	CP014658.1	<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar Anatum	USDA-ARS-USMARC-1736	286
	pEC732_IMP14	IncC	CP015139.1	<i>Escherichia coli</i>	Ecol_732	286
	pHS36-NDM	IncC	KU726616.1	<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar Stanley	LS001	286

Cluster	Plasmide	Groupe Inc	Accession	Espèce	Souche	Longueur (aa)
	pGPN1	IncC	CP011540.1	<i>Klebsiella aerogenes</i>	G7	286
	pVAS3-1	IncC	KU160531.1	<i>Vibrio alginolyticus</i>	VAS3-1	286
	pV004-a	IncC	LC056141.1	<i>Escherichia coli</i>	V004_1	286
	pV001-a	IncC	LC056079.1	<i>Escherichia coli</i>	V001_1	286
	pV266-a	IncC	LC056485.1	<i>Escherichia coli</i>	V266_1	286
	pEcl-Gr4873	IncC	KR559890.1	<i>Enterobacter cloacae</i>	Ecl4873	286
	pKP-Gr8143	IncC	KR559889.1	<i>Klebsiella pneumoniae</i>	Kpn8143	286
	pKP-Gr642	IncC	KR559888.1	<i>Klebsiella pneumoniae</i>	Kpn642	286
	pNDM15-1078	IncC	CP012902.1	<i>Escherichia coli</i>	N15-01078	286
	pCAV1193-166	IncC	CP013324.1	<i>Klebsiella pneumoniae</i>	CAV1193	286
	pRH-1238	IncC	KR091911.1	<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar Corvallis		286
	pKPC_CAV1344	IncC	CP011622.1	<i>Klebsiella pneumoniae</i>	CAV1344	286
	pMRVIM0912	IncC	KP975074.1	<i>Citrobacter freundii</i>	MRSN11938	286
	pNDM-116-17_1	IncC	LN831185.1	<i>Vibrio cholerae</i>		286
	pHM881QN	IncC	LC055503.1	<i>Klebsiella pneumoniae</i>	Y881_1	286
	pNDM-SAL	IncC	KP742988.1	<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar Senftenberg	BCH02406	286
	pYU39_IncA_C	IncC	CP011429.1	<i>Salmonella enterica</i> subsp. <i>enterica</i>	YU39	286
	pNDM-PstGN576	IncC	KJ802405.1	<i>Providencia stuartii</i>	GN576	286
	pNDM-EcoGN568	IncC	KJ802404.1	<i>Escherichia coli</i>	GN568	286
	pBK32533	IncC	KP345882.1	<i>Escherichia coli</i>	BK32533	286
	pVC1307	IncC	KJ817377.1	<i>Vibrio cholerae</i> O1	ICDC-1307	286
	pCFSAN007428_01	IncC	CP009414.2	<i>Salmonella enterica</i>	CFSAN007428	286

Cluster	Plasmide	Groupe Inc	Accession	Espèce	Souche	Longueur (aa)
	pCFSAN007425_01	IncC	CP009411.2	<i>Salmonella enterica</i>	CFSAN007425	286
	pCFSAN000405_01	IncC	CP009409.2	<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar Heidelberg	N418	286
	p39R861-4	IncC	KP276584.1	<i>Escherichia coli</i>	39R861	286
	p6409-202.186kb	IncC	CP010373.2	<i>Escherichia coli</i>	6409	286
	pNDM-US-2	IncC	KJ588779.1	<i>Klebsiella pneumoniae</i>	ATCC_B-2146	286
	p11663	IncC	AP014611.1	<i>Serratia marcescens</i>	11663_1	286
	p6234-178.193kb	IncC	CP010391.1	<i>Klebsiella pneumoniae</i>	6234	286
	pVC1447	IncC	KM083064.1	<i>Vibrio cholerae</i>	ICDC-1447	286
	pYDC637	IncC	KP056256.1	<i>Escherichia coli</i>	YDC637	286
	pPSP-100	IncC	CP009868.1	<i>Pantoea sp.</i>	PSNIH2	286
	pSRC119-A_C	IncC	KM670336.1	<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar Senftenberg	SRC119	286
	pCVM22425	IncC	CP009560.1	<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar Newport	CVM_22425	286
	pCFSAN000934_02	IncC	CP009567.1	<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar Newport	CVM_22462	286
	pVch2012EL-2176	IncC	CP007636.1	<i>Vibrio cholerae</i>	2012EL-2176	286
	pSN254b	IncC	KJ909290.1	<i>Aeromonas salmonicida</i> subsp. <i>salmonicida</i>	2004-05MF26	286
	Chromosomique	NA	AP014565.1	<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar Typhimurium	L-3553	286
	pKOX-86d	IncC	CP008790.1	<i>Klebsiella oxytoca</i>	KONIH1	286

Cluster	Plasmide	Groupe Inc	Accession	Espèce	Souche	Longueur (aa)
	p1643_10	IncC	KF056330.2	<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar Kentucky	1643_10	286
	pRMH760	IncC	KF976462.2	<i>Klebsiella pneumoniae</i>		286
	pNDM-US	IncC	CP006661.1	<i>Klebsiella pneumoniae</i>	ATCC_B-2146	286
	pVCR94deltaX	IncC	KF551948.1	<i>Vibrio cholerae</i>	BI144	286
	pSCEC2	IncC	KF152885.1	<i>Escherichia coli</i>	SCEC2	286
	pIncA/C-LS6	IncC	JX442976.1	<i>Klebsiella pneumoniae</i>		286
	pEA1509_A	IncC	FO203354.1	<i>Enterobacter aerogenes</i>	EA1509E	286
	pNDM102337	IncC	JF714412.2	<i>Escherichia coli</i>	N10-2337	286
	pR148	IncC	JX141473.1	<i>Aeromonas hydrophila</i>		286
	pTC2	IncC	JQ824049.1	<i>Providencia stuartii</i>		286
	pSH111_166	IncC	JN983043.1	<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar Heidelberg		286
	pNDM10469	IncC	JN861072.1	<i>Klebsiella pneumoniae</i>	N10-0469	286
	pR55	IncC	JQ010984.1	<i>Klebsiella pneumoniae</i>		286
	pMR0211	IncC	JN687470.1	<i>Providencia stuartii</i>		286
	pNDM-KN	IncC	JN157804.1	<i>Klebsiella pneumoniae</i>	Kp7	286
	pNDM10505	IncC	JF503991.1	<i>Escherichia coli</i>	N10-0505	286
	pNDM-1_Dok01	IncC	AP012208.1	<i>Escherichia coli</i>	NDM-1_Dok01_1	286
	pUMNK88	IncC	HQ023862.1	<i>Escherichia coli</i>	UMNK88	286
	pPG010208	IncC	HQ023861.1	<i>Escherichia coli</i>	PG010208	286
	Chromosomique	NA	AB571791.1	<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar Typhimurium		286
	pSD_174	IncC	JF267651.1	<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar Dublin	853	286

Cluster	Plasmide	Groupe Inc	Accession	Espèce	Souche	Longueur (aa)
	pAR060302	IncC	FJ621588.1	<i>Escherichia coli</i>	AR060302	286
	pAM04528	IncC	FJ621587.1	<i>Salmonella enterica</i>	AM04528	286
	pIP1202	IncC	CP000603.1	<i>Yersinia pestis</i> biovar <i>Orientalis</i>	IP275	286
	pSN254	IncC	CP000604.1	<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar Newport	SL254	286
	pYR1	IncC	CP000602.1	<i>Yersinia ruckeri</i>	YR71	286
	pP91278	IncC	AB277724.1	<i>Photobacterium damsela</i> subsp. <i>piscicida</i>		286
	pP99-018	IncC	AB277723.1	<i>Photobacterium damsela</i> subsp. <i>piscicida</i>		286
	pGMI14-002_1	IncC	CP028197.1	<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar Concord	CFSAN018747	286
	pAPEC1990_61	IncC	HQ023863.1	<i>Escherichia coli</i>	APEC1990_61	286
	pCFSAN007427_01	IncC	CP009413.2	<i>Salmonella enterica</i>	CFSAN007427	270
	p1761_02	IncC	CP039976.1	<i>Klebsiella pneumoniae</i>	R1761	286
pAQU1	pAQU1	Untyped	AB571865.1	<i>Photobacterium damsela</i> subsp. <i>damsela</i>	04Ya311_1	272
	pVPS62	Untyped	KX957971.1	<i>Vibrio parahaemolyticus</i>		272
	p345-185	Untyped	CP025539.1	<i>Vibrio harveyi</i>	345	272
	pVPS43	Untyped	KX957970.1	<i>Vibrio parahaemolyticus</i>		272
	pVPH2	Untyped	KP791968.1	<i>Vibrio parahaemolyticus</i>	2011VPH2	272
	pVAS114	Untyped	KX957969.1	<i>Vibrio alginolyticus</i>		272

Cluster	Plasmide	Groupe Inc	Accession	Espèce	Souche	Longueur (aa)
	pATCC_33868-1	Untyped	CP014052.1	<i>Vibrio alginolyticus</i>	ARGOS_108	272
	pAQU2	Untyped	AB856327.1	<i>Vibrio sp. 04Ya090</i>		272
pIncAC-KP4898	pIncAC-KP4898	IncA	KY882285.1	<i>Klebsiella pneumoniae</i> subsp. <i>pneumoniae</i>	KP4898	286
	pIBAC_IncA/C	IncA	MH594477.1	<i>Citrobacter freundii</i>	535	286
	pKC-BO-N1-VIM	IncA	MG228427.1	<i>Kluyvera cryocrescens</i>	BO64W	286
	pIBAC_Incx3_A/C	IncA/IncX3	MH594478.1	<i>Citrobacter freundii</i>	593	286
	p34998-239.973kb	IncA	CP012168.1	<i>Enterobacter hormaechei</i> subsp. <i>steigerwaltii</i>	34998	286
	pR210-2-VIM	IncA	CP034084.1	<i>Klebsiella pneumoniae</i> subsp. <i>pneumoniae</i>	R210-2	286
pRA1	pRA1	IncA	FJ705807.1	<i>Aeromonas hydrophila</i>		286
	pEc09	IncA	MH325468.1	<i>Enterobacter hormaechei</i>	Ec09	286
	pG3216.2	IncA	MG550958.1	<i>Escherichia coli</i>	G3216	286
	pCf587	IncA	MG053108.1	<i>Citrobacter freundii</i>	33587	286
	pEc13	IncA	MH325469.1	<i>Enterobacter hormaechei</i>	Ec13	286
	pRSE15	IncA	CP034714.1	<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar Mikawasima	RSE15	286
PmCHE-GI1	Chromosomique	NA	KJ439039.1	<i>Proteus mirabilis</i>	PmCHE	286
	Chromosomique	NA	KF856624.1	<i>Proteus mirabilis</i>	PEL	286
	Chromosomique	NA	KJ411925.1	<i>Proteus mirabilis</i>	PmCHA	286
	Chromosomique	NA	CP020052.1	<i>Proteus mirabilis</i>	AR_0059	286
	Chromosomique	NA	CP001120.1	<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar Heidelberg	SL476	286

Cluster	Plasmide	Groupe Inc	Accession	Espèce	Souche	Longueur (aa)
	Chromosomique	NA	KU499917.1	<i>Proteus mirabilis</i>	PmESC	286
pHDAS3.2	pHDAS3.2	Untyped	CP017612.1	<i>Candidatus Hamiltonella defensa</i>	AS3	283
	pHDAS3.2	Untyped	CP017612.1	<i>Candidatus Hamiltonella defensa</i>	AS3	205
	pHDA2C.2	Untyped	CP017608.1	<i>Candidatus Hamiltonella defensa</i>	A2C	283
	pHDA2C.2	Untyped	CP017608.1	<i>Candidatus Hamiltonella defensa</i>	A2C	283
pSTY1-2010K-1587	pSTY1-2010K-1587	IncC	CP016864.1	<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar Typhimurium	CDC_2010K-1587	286
	pCFSAN001921-1	IncC	CP006050.1	<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar Typhimurium var. 5	CFSAN001921	286
	pSenHei5-3	IncC	CP031362.1	<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar Heidelberg	5	286
pKHM-1	pKHM-1	Untyped	AP014939.1	<i>Citrobacter freundii</i>	KHM_243_1	287
	pFERM_P-14531-1	Untyped	AP018687.1	<i>Vibrio rumoiensis</i>	FERM_P-14531	287
pXNC1	pXNC1	Untyped	FN667743.1	<i>Xenorhabdus nematophila</i>	ATCC_19061	287
	pXNC2	Untyped	LN681228.1	<i>Xenorhabdus nematophila</i>	AN6_1	287
pIP40a	pIP40a	IncC	KX709966.1	<i>Pseudomonas aeruginosa</i>	IP40a	286
	IP40a	IncC	KX156772.1	<i>Escherichia coli</i>	K-12	286
pKEC-a3c	pKEC-a3c	IncC	CP007558.1	<i>Citrobacter freundii</i>	CFNIH1	286

Cluster	Plasmide	Groupe Inc	Accession	Espèce	Souche	Longueur (aa)
	pKEC-dc3	IncC	CP007732.1	<i>Klebsiella pneumoniae</i> subsp. <i>pneumoniae</i>	KPNIH27	286
pAsa4c	pAsa4c	Untyped	KT033470.1	<i>Aeromonas salmonicida</i> subsp. <i>salmonicida</i>	JF2267	286
	pS44-1	Untyped	CP022176.1	<i>Aeromonas salmonicida</i>	S44	286
pEan1	pEan1	Untyped	CP006665.1	<i>Edwardsiella anguillarum</i>	ET080813	283
	pEP-MS-18-199	Untyped	CP035669.1	<i>Edwardsiella piscicida</i>	MS-18-199	283
<i>V. parahaemolyticus</i> 10329	Chromosomique	NA	CP026041.1	<i>Vibrio parahaemolyticus</i>	10329	273
	Chromosomique	NA	CP023248.1	<i>Vibrio parahaemolyticus</i>	MAVP-26	273
<i>A. veronii</i> CB51 (1)	Chromosomique	NA	CP015448.1	<i>Aeromonas veronii</i>	CB51	289
pIMP-PH114	pIMP-PH114	IncC	KF250428.1	<i>Klebsiella pneumoniae</i>	CRE114	286
pDGO100	pDGO100	IncC	KU997026.1	<i>Escherichia coli</i>	VA292	286
pCf164_LMB-1	pCf164_LMB-1	IncC	MH475146.1	<i>Citrobacter freundii</i>	164	286
pVPS129	pVPS129	IncC	KY014464.1	<i>Vibrio parahaemolyticus</i>		286
pAsa4b	pAsa4b	Untyped	KT033469.1	<i>Aeromonas salmonicida</i> subsp. <i>salmonicida</i>	01-B522	286
pAhD4-1	pAhD4-1	Untyped	CP013966.1	<i>Aeromonas hydrophila</i>	D4	286
<i>V. cholerae</i> 10432-62	Chromosomique	NA	CP010812.1	<i>Vibrio cholerae</i>	10432-62	285
<i>A. veronii</i> CB51 (2)	Chromosomique	NA	CP015448.1	<i>Aeromonas veronii</i>	CB51	283
<i>V. campbellii</i> 20130629003 S01 (1)	Chromosomique	NA	CP020076.1	<i>Vibrio campbellii</i>	20130629003S01	282

Cluster	Plasmide	Groupe Inc	Accession	Espèce	Souche	Longueur (aa)
<i>V. diabolicus</i> FDRGOS_96	Chromosomique	NA	CP014094.1	<i>Vibrio diabolicus</i>	FDRGOS_96	282
<i>V. parahaemolyticus</i> S107-1	Chromosomique	NA	CP028481.1	<i>Vibrio parahaemolyticus</i>	S107-1	276
<i>V. anguillarum</i> CNEVA_NB11008	Chromosomique	NA	CP022103.1	<i>Vibrio anguillarum</i>	CNEVA_NB11008	276
<i>V. parahaemolyticus</i> D3112	Chromosomique	NA	CP034565.1	<i>Vibrio parahaemolyticus</i>	D3112	276
<i>V. vulnificus</i> 93U204	Chromosomique	NA	CP009261.1	<i>Vibrio vulnificus</i>	93U204	276
<i>V. campbellii</i> 20130629003S01 (2)	Chromosomique	NA	CP020076.1	<i>Vibrio campbellii</i>	20130629003S01	276
<i>V. harveyi</i> QT520	Chromosomique	NA	CP018680.2	<i>Vibrio harveyi</i>	QT520	276
<i>V. campbellii</i> BoB-90	Chromosomique	NA	CP026315.1	<i>Vibrio campbellii</i>	BoB-90	276
<i>V. parahaemolyticus</i> FORC_004	Chromosomique	NA	CP009847.1	<i>Vibrio parahaemolyticus</i>	FORC_004	276
<i>V. parahaemolyticus</i> FORC_023	Chromosomique	NA	CP012950.1	<i>Vibrio parahaemolyticus</i>	FORC_023	276
<i>V. neocaledonicus</i> GCJ02-2	Chromosomique	NA	CP032213.1	<i>Vibrio neocaledonicus</i>	CGJ02-2	276

Cluster	Plasmide	Groupe Inc	Accession	Espèce	Souche	Longueur (aa)
<i>V. cholerae</i> FORC_055	Chromosomique	NA	CP016987.1	<i>Vibrio cholerae</i>	FORC_055	275
<i>V. parahaemolyticus</i> S107-1	Chromosomique	NA	CP028481.1	<i>Vibrio parahaemolyticus</i>	S107-1	275
pPagHao2018	pPagHao2018	Untyped	CP033067.1	<i>Pseudoalteromonas agarivorans</i>	Hao_2018	273
pVPS91	pVPS91	Untyped	KX957972.1	<i>Vibrio parahaemolyticus</i>		272
pVAS19	pVAS19	Untyped	KX957968.1	<i>Vibrio alginolyticus</i>		272
pVPH1	pVPH1	Untyped	KP688397.1	<i>Vibrio parahaemolyticus</i>	V36	272
<i>X. poinarii</i> G6	Chromosomique	NA	FO704551.1	<i>Xenorhabdus poinarii</i>	G6	261
<i>P. stutzeri</i> 19SMN4	Chromosomique	NA	CP007509.1	<i>Pseudomonas stutzeri</i>	19SMN4	257
<i>A. caviae</i> 8LM	Chromosomique	NA	CP024198.1	<i>Aeromonas caviae</i>	8LM	256

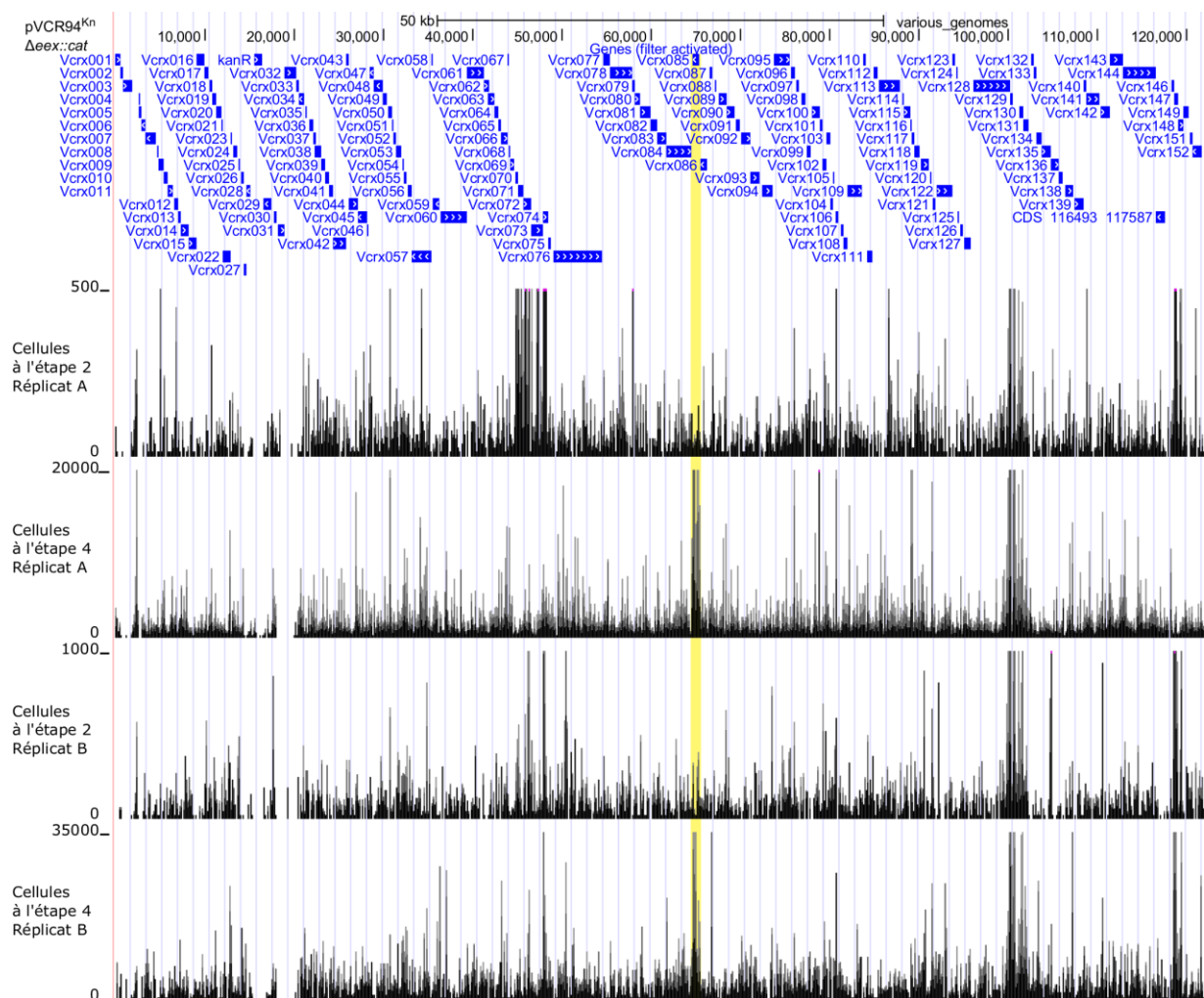


Figure S1. Identification par HDTM d'un potentiel déterminant de l'exclusion, résultats complets

Densité d'insertion de transposons dans pVCR94^{Kn}Δ*ex::cat* avant (étape 2) et après (étape 4) transfert de pClo. La région correspondant au gène *vcrx085* est surlignée en jaune

BIBLIOGRAPHIE

- Achtman, M., Kennedy, N., and Skurray, R. (1977). Cell--cell interactions in conjugating *Escherichia coli*: role of traT protein in surface exclusion. *Proc Natl Acad Sci U S A* 74, 5104–5108.
- Almagro Armenteros, J.J., Tsirigos, K.D., Sønderby, C.K., Petersen, T.N., Winther, O., Brunak, S., von Heijne, G., and Nielsen, H. (2019). SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat. Biotechnol.* 37, 420–423.
- Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410.
- Ambrose, S.J., Harmer, C.J., and Hall, R.M. (2018). Compatibility and entry exclusion of IncA and IncC plasmids revisited: IncA and IncC plasmids are compatible. *Plasmid* 96–97, 7–12.
- Anthony, K.G., Sherburne, C., Sherburne, R., and Frost, L.S. (1994). The role of the pilus in recipient cell recognition during bacterial conjugation mediated by F-like plasmids. *Mol. Microbiol.* 13, 939–953.
- Anthony, K.G., Klimke, W.A., Manchak, J., and Frost, L.S. (1999). Comparison of proteins involved in pilus synthesis and mating pair stabilization from the related plasmids F and R100-1: insights into the mechanism of conjugation. *J Bacteriol* 181, 5149–5159.
- Audette, G.F., Manchak, J., Beatty, P., Klimke, W.A., and Frost, L.S. (2007). Entry exclusion in F-like plasmids requires intact TraG in the donor that recognizes its cognate TraS in the recipient. *Microbiology* 153, 442–451.
- Babić, A., Lindner, A.B., Vulić, M., Stewart, E.J., and Radman, M. (2008). Direct visualization of horizontal gene transfer. *Science* 319, 1533–1536.
- Beaber, J.W., Hochhut, B., and Waldor, M.K. (2004). SOS response promotes horizontal dissemination of antibiotic resistance genes. *Nature* 427, 72–74.
- Burrus, V. (2017). Mechanisms of stabilization of integrative and conjugative elements. *Current Opinion in Microbiology* 38, 44–50.
- Cabezón, E., Sastre, J.I., and de la Cruz, F. (1997). Genetic evidence of a coupling role for the TraG protein family in bacterial conjugation. *Mol. Gen. Genet.* 254, 400–406.
- Cabezón, E., Ripoll-Rozada, J., Peña, A., Cruz, F. de la, and Arechaga, I. (2015). Towards an integrated model of bacterial conjugation. *FEMS Microbiology Reviews* 39, 81–95.
- Capella-Gutiérrez, S., Silla-Martínez, J.M., and Gabaldón, T. (2009). trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 25, 1972–1973.

Carraro, N., Matteau, D., Luo, P., Rodrigue, S., and Burrus, V. (2014a). The master activator of IncA/C conjugative plasmids stimulates genomic islands and multidrug resistance dissemination. *PLoS Genet* 10.

Carraro, N., Sauv  , M., Matteau, D., Lauzon, G., Rodrigue, S., and Burrus, V. (2014b). Development of pVCR94  X from *Vibrio cholerae*, a prototype for studying multidrug resistant IncA/C conjugative plasmids. *Front Microbiol* 5.

Carraro, N., Matteau, D., Burrus, V., and Rodrigue, S. (2015). Unraveling the regulatory network of IncA/C plasmid mobilization: When genomic islands hijack conjugative elements. *Mob Genet Elements* 5, 34–38.

Carraro, N., Rivard, N., Ceccarelli, D., Colwell, R.R., and Burrus, V. (2016). IncA/C conjugative plasmids mobilize a new family of multidrug resistance islands in clinical *Vibrio cholerae* non-O1/non-O139 isolates from Haiti. *MBio* 7.

Carraro, N., Durand, R., Rivard, N., Anquetil, C., Barrette, C., Humbert, M., and Burrus, V. (2017a). *Salmonella* genomic island 1 (SGI1) reshapes the mating apparatus of IncC conjugative plasmids to promote self-propagation. *PLOS Genetics* 13, e1006705.

Carraro, N., Rivard, N., Burrus, V., and Ceccarelli, D. (2017b). Mobilizable genomic islands, different strategies for the dissemination of multidrug resistance and other adaptive traits. *Mob Genet Elements* 7, 1–6.

Ceccarelli, D., Daccord, A., Ren  , M., and Burrus, V. (2008). Identification of the origin of transfer (*oriT*) and a new gene required for mobilization of the SXT/R391 family of integrating conjugative elements. *J. Bacteriol.* 190, 5328–5338.

Cherepanov, P.P., and Wackernagel, W. (1995). Gene disruption in *Escherichia coli*: Tc^R and Km^R cassettes with the option of FLP-catalyzed excision of the antibiotic-resistance determinant. *Gene* 158, 9–14.

Christie, P.J. (2004). Type IV secretion: the *Agrobacterium* VirB/D4 and related conjugation systems. *Biochim Biophys Acta* 1694, 219–234.

Christie, P.J. (2016). The mosaic type IV secretion systems. *EcoSal Plus* 7.

Clarke, M., Maddera, L., Harris, R.L., and Silverman, P.M. (2008). F-pili dynamics by live-cell imaging. *Proc Natl Acad Sci U S A* 105, 17978–17981.

Cruz, F.D.L., Frost, L.S., Meyer, R.J., and Zechner, E.L. (2010). Conjugative DNA metabolism in Gram-negative bacteria. *FEMS Microbiology Reviews* 34, 18–40.

Datsenko, K.A., and Wanner, B.L. (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A* 97, 6640–6645.

Datta, S., Costantino, N., and Court, D.L. (2006). A set of recombineering plasmids for gram-negative bacteria. *Gene* 379, 109–115.

Doublet, B., Golding, G.R., Mulvey, M.R., and Cloeckaert, A. (2008). Secondary chromosomal attachment site and tandem integration of the mobilizable *Salmonella* Genomic Island 1. *PLoS One* 3.

Dunny, G.M., and Hirt, H. (2019). A new flavor of entry exclusion in ICE elements provides a selective advantage for the element and its host. *Molecular Microbiology* 112, 1061–1065.

Dürrenberger, M.B., Villiger, W., and Bächli, Th. (1991). Conjugational junctions: Morphology of specific contacts in conjugating *Escherichia coli* bacteria. *Journal of Structural Biology* 107, 146–156.

Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797.

Finnegan, J., and Sherratt, D. (1982). Plasmid ColE1 conjugal mobility: the nature of *bom*, a region required in *cis* for transfer. *Mol. Gen. Genet.* 185, 344–351.

Gago-Córdoba, C., Val-Calvo, J., Miguel-Arribas, A., Serrano, E., Singh, P.K., Abia, D., Wu, L.J., and Meijer, W.J.J. (2019). Surface exclusion revisited: function related to differential expression of the surface exclusion system of *Bacillus subtilis* plasmid pLS20. *Front Microbiol* 10.

Garcillán-Barcia, M.P., and de la Cruz, F. (2008). Why is entry exclusion an essential feature of conjugative plasmids? *Plasmid* 60, 1–18.

Gardy, J.L., Spencer, C., Wang, K., Ester, M., Tusnády, G.E., Simon, I., Hua, S., deFays, K., Lambert, C., Nakai, K., et al. (2003). PSORT-B: improving protein subcellular localization prediction for Gram-negative bacteria. *Nucleic Acids Res* 31, 3613–3617.

Grenier, F., Matteau, D., Baby, V., and Rodrigue, S. (2014). Complete genome sequence of *Escherichia coli* BW25113. *Genome Announc* 2, pii: e01038-14.

Grohmann, E., Christie, P.J., Waksman, G., and Backert, S. (2018). Type IV secretion in Gram-negative and Gram-positive bacteria. *Molecular Microbiology* 107, 455–471.

Guglielmini, J., Cruz, F. de la, and Rocha, E.P.C. (2013). Evolution of conjugation and type IV secretion systems. *Mol Biol Evol* 30, 315–331.

Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* 59, 307–321.

Gunton, J.E., Ussher, J.E.R., Rooker, M.M., Wetsch, N.M., Alonso, G., and Taylor, D.E. (2008). Entry exclusion in the IncHI1 plasmid R27 is mediated by EexA and EexB. *Plasmid* 59, 86–101.

Guzman, L.M., Belin, D., Carson, M.J., and Beckwith, J. (1995). Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *J Bacteriol* 177, 4121–4130.

Haase, J., Kalkum, M., and Lanka, E. (1996). TrbK, a small cytoplasmic membrane lipoprotein, functions in entry exclusion of the IncP alpha plasmid RP4. *J Bacteriol* 178, 6720–6729.

Hancock, S.J., Phan, M.-D., Peters, K.M., Forde, B.M., Chong, T.M., Yin, W.-F., Chan, K.-G., Paterson, D.L., Walsh, T.R., Beatson, S.A., et al. (2017). Identification of IncA/C plasmid replication and maintenance genes and development of a plasmid multilocus sequence typing scheme. *Antimicrob Agents Chemother* 61.

Hancock, S.J., Phan, M.-D., Luo, Z., Lo, A.W., Peters, K.M., Nhu, N.T.K., Forde, B.M., Whitfield, J., Yang, J., Strugnelli, R.A., et al. (2020). Comprehensive analysis of IncC plasmid conjugation identifies a crucial role for the transcriptional regulator AcaB. *Nature Microbiology* 1–9.

Harmer, C.J., and Hall, R.M. (2015). The A to Z of A/C plasmids. *Plasmid* 80, 63–82.

Harmer, C.J., Hamidian, M., Ambrose, S.J., and Hall, R.M. (2016). Destabilization of IncA and IncC plasmids by SGI1 and SGI2 type *Salmonella* genomic islands. *Plasmid*.

Hartskeerl, R., Overduin, P., Hoekstra, W., and Tommassen, J. (1986). Nucleotide sequence of the exclusion-determining locus of IncI plasmid R144. *Gene* 42, 107–111.

Hegy, A., Szabó, M., Olasz, F., and Kiss, J. (2017). Identification of *oriT* and a recombination hot spot in the IncA/C plasmid backbone. *Scientific Reports* 7, 10595.

Hu, B., Khara, P., and Christie, P.J. (2019). Structural bases for F plasmid conjugation and F pilus biogenesis in *Escherichia coli*. *Proc Natl Acad Sci U S A* 116, 14222–14227.

Huguet, K.T., Rivard, N., Garneau, D., Palanee, J., and Burrus, V. (2020). Replication of the *Salmonella* Genomic Island 1 (SGI1) triggered by helper IncC conjugative plasmids promotes incompatibility and plasmid loss. *PLOS Genetics* 16, e1008965.

Humbert, M., Huguet, K.T., Coulombe, F., and Burrus, V. (2019). Entry exclusion of conjugative plasmids of the IncA, IncC, and related untyped incompatibility groups. *Journal of Bacteriology* 201, e00731-18, /jb/201/10/JB.00731-18.atom.

Jalajakumari, M.B., Guidolin, A., Buhk, H.J., Manning, P.A., Ham, L.M., Hodgson, A.L.M., Cheah, K.C., and Skurray, R.A. (1987). Surface exclusion genes *traS* and *traT* of the F sex factor of *Escherichia coli* K-12: Determination of the nucleotide sequence and promoter and terminator activities. *Journal of Molecular Biology* 198, 1–11.

Johnson, C.M., and Grossman, A.D. (2015). Integrative and conjugative elements (ICEs): What they do and how they work. *Annu Rev Genet* 49, 577–601.

Jones, D.T. (1999). Protein secondary structure prediction based on position-specific scoring matrices. *J. Mol. Biol.* 292, 195–202.

Klimke, W.A., and Frost, L.S. (1998). Genetic analysis of the role of the transfer gene, *traN*, of the F and R100-1 plasmids in mating pair stabilization during conjugation. *J Bacteriol* 180, 4036–4043.

- Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 33, 1870–1874.
- Lawley, T.D., Klimke, W.A., Gubbins, M.J., and Frost, L.S. (2003). F factor conjugation is a true type IV secretion system. *FEMS Microbiology Letters* 224, 1–15.
- Lederberg, J., Cavalli, L.L., and Lederberg, E.M. (1952). Sex compatibility in *Escherichia coli*. *Genetics* 37, 720–730.
- Li, W., and Godzik, A. (2006). Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* 22, 1658–1659.
- Madeira, F., Park, Y., Lee, J., Buso, N., Gur, T., Madhusoodanan, N., Basutkar, P., Tivey, A., Potter, S., Finn, R., et al. (2019). The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res* 47, W636–W641.
- Marrero, J. (2007). Entry exclusion in ICEs: defining the role of the recipient during the conjugative process. Tufts University - Sackler School of Graduate Biomedical Sciences 158.
- Marrero, J., and Waldor, M.K. (2005). Interactions between inner membrane proteins in donor and recipient cells limit conjugal DNA transfer. *Dev. Cell* 8, 963–970.
- Marrero, J., and Waldor, M.K. (2007a). The SXT/R391 family of integrative conjugative elements is composed of two exclusion groups. *J Bacteriol* 189, 3302–3305.
- Marrero, J., and Waldor, M.K. (2007b). Determinants of entry exclusion within Eex and TraG are cytoplasmic. *J Bacteriol* 189, 6469–6473.
- Moll, A., Manning, P.A., and Timmis, K.N. (1980). Plasmid-determined resistance to serum bactericidal activity: a major outer membrane protein, the *traT* gene product, is responsible for plasmid-specified serum resistance in *Escherichia coli*. *Infect Immun* 28, 359–367.
- O'Neill, J. (2016). Tackling drug-resistant infections globally: final report and recommendations.
- Patel, S. (2016). Drivers of bacterial genomes plasticity and roles they play in pathogen virulence, persistence and drug resistance. *Infection, Genetics and Evolution* 45, 151–164.
- Pohlman, R.F., Genetti, H.D., and Winans, S.C. (1994). Entry exclusion of the IncN plasmid pKM101 is mediated by a single hydrophilic protein containing a lipid attachment motif. *Plasmid* 31, 158–165.
- Rivard, N., Colwell, R.R., and Burrus, V. (2020). Antibiotic resistance in *Vibrio cholerae*: Mechanistic insights from IncC plasmid-mediated dissemination of a novel family of genomic islands inserted at *trmE*. *MSphere* 5.
- Roy, D., Huguet, K.T., Grenier, F., and Burrus, V. (2020). IncC conjugative plasmids and SXT/R391 elements repair double-strand breaks caused by CRISPR–Cas during conjugation. *Nucleic Acids Res.*

Rozwandowicz, M., Brouwer, M.S.M., Fischer, J., Wagenaar, J.A., Gonzalez-Zorn, B., Guerra, B., Mevius, D.J., and Hordijk, J. (2018). Plasmids carrying antimicrobial resistance genes in Enterobacteriaceae. *J Antimicrob Chemother* 73, 1121–1137.

Sakuma, T., Tazumi, S., Furuya, N., and Komano, T. (2013). ExcA proteins of IncI1 plasmid R64 and IncIy plasmid R621a recognize different segments of their cognate TraY proteins in entry exclusion. *Plasmid* 69, 138–145.

Samuels, A.L., Lanka, E., and Davies, J.E. (2000). Conjugative junctions in RP4-mediated mating of *Escherichia coli*. *J Bacteriol* 182, 2709–2715.

Sheehy, R.J., Orr, C., and Curtiss, R. (1972). Molecular studies on entry exclusion in *Escherichia coli* minicells. *J Bacteriol* 112, 861–869.

Singer, M., Baker, T.A., Schnitzler, G., Deischel, S.M., Goel, M., Dove, W., Jaacks, K.J., Grossman, A.D., Erickson, J.W., and Gross, C.A. (1989). A collection of strains containing genetically linked alternating antibiotic resistance elements for genetic mapping of *Escherichia coli*. *Microbiol. Rev.* 53, 1–24.

Soucy, S.M., Huang, J., and Gogarten, J.P. (2015). Horizontal gene transfer: building the web of life. *Nature Reviews Genetics* 16, 472–482.

Thomas, C.M., Thomson, N.R., Cerdeño-Tárraga, A.M., Brown, C.J., Top, E.M., and Frost, L.S. (2017). Annotation of plasmid genes. *Plasmid* 91, 61–67.

Waksman, G. (2019). From conjugation to T4S systems in Gram-negative bacteria: a mechanistic biology perspective. *EMBO Rep* 20.

Zechner, E.L., Lang, S., and Schildbach, J.F. (2012). Assembly and mechanisms of bacterial type IV secretion machines. *Philos Trans R Soc Lond B Biol Sci* 367, 1073–1087.